Exploring Links between Mother and Infant through Human Milk
Diurnal Melatonin Rhythm in Human Milk and Association of Maternal Diet and BMI with Infant Gut Bacteria

Master’s Thesis
In partial fulfilment of the requirements for the Master of Science ETH in Pharmacy

Submitted by:
Viktoria Gastens

Supervisor:
PD Dr. Vivianne Otto
Institute of Pharmaceutical Sciences / ETH Zurich

Co-supervisors:
Dr. Salome Kurth
Sarah Schoch
Baby Sleep Laboratory / CRPP Sleep and Health / University Hospital Zurich

Zurich, July 2018
Contents

1 Abstract 3

2 Introduction 4
   2.1 Motivation . 4
   2.2 Human milk 5
      2.2.1 Impact of breastfeeding 5
      2.2.2 Incidence of breastfeeding 7
      2.2.3 Anatomy and physiology of lactation 7
      2.2.4 Composition of human milk 9
      2.2.5 Human milk microbiome 9
      2.2.6 Human milk and sleep 11
   2.3 Melatonin 11
      2.3.1 Impact of Melatonin 11
      2.3.2 Physiology of Melatonin 12
   2.4 Restaurant hypothesis 13
   2.5 Gut microbiota 14
      2.5.1 Impact of the early gut microbiota 14
      2.5.2 Development of the early gut microbiota 15
      2.5.3 Diurnal rhythmicity of gut microbiota 16
   2.6 Diet and gut microbiota 16
      2.6.1 Diet and gut microbiota in adults 16
      2.6.2 Maternal diet and infant gut microbiota 17
   2.7 Body Mass Index 17
      2.7.1 Body Mass Index and gut microbiota 18
   2.8 Guidelines 18
      2.8.1 Breastfeeding practices 18
      2.8.2 Nutrition during the lactation period 19

3 Aim 20
# Methods

4.1 Pilot Experiment
   4.1.1 Recruiting
   4.1.2 Home visit
   4.1.3 Sampling

4.2 Human milk analysis
   4.2.1 Sample preparation
   4.2.2 ELISA analysis of Melatonin

4.3 Infant gut bacteria study
   4.3.1 Participants
   4.3.2 Process

4.4 Longitudinal study and further involvement

4.5 Analysis of infant gut bacteria
   4.5.1 16S rRNA gene analysis
   4.5.2 Data analysis

4.6 Statistical analysis

# Results

5.1 Pilot Experiment
   5.1.1 Sampling
   5.1.2 ELISA analysis of Melatonin

5.2 Maternal diet, Body Mass Index and infant gut bacteria
   5.2.1 Infant gut bacteria at 3 months
   5.2.2 Food Frequency Questionnaire distributions
   5.2.3 Body Mass Index distribution
   5.2.4 Association of maternal diet and infant gut bacteria
   5.2.5 Regression of infant gut bacteria and maternal BMI

# Discussion

6.1 Diurnal Melatonin rhythm in human milk

6.2 Association of maternal diet and BMI with infant gut bacteria

6.3 Conclusions

6.4 Outlook

# Acknowledgements

References

Appendices
Chapter 1

Abstract

A comprehensive overview of the scientific efforts towards understanding the beneficial effects of breastfeeding on infants, mothers and societies has been collected. Of particular interest are the mechanisms which change the human milk composition dynamically over the lactation period, day and feed. Two potential links between mother and infant through human milk were investigated: diurnal Melatonin rhythm in human milk and association of maternal diet and BMI with infant gut bacteria.

A pilot study was designed to find a feasible way of monitoring human milk Melatonin levels. I decided to use a commercial ELISA kit designed for human serum and plasma. The human milk Melatonin concentration curve peaked at 11 pg/mL between 2 and 5 am. During daytime, the Melatonin level dropped below the ELISA detection limit (1.6 pg/mL). The diurnal rhythm was clearly detectable, making this procedure feasible for the purpose of correlating human milk Melatonin and infant sleep behaviour.

In the longitudinal study *Sleep Behavior and Gut Microbiota in Healthy Infants*, maternal diet was monitored with a self-reported Food Frequency Questionnaire and infant gut bacteria were analysed by 16S rRNA gene analysis. I was not able to find any correlation between maternal food intake and their breastfed infants’ gut bacteria. There was a small effect of maternal BMI on *Actinobacteria* ($r_{Pearson} = .18$) and the *Bacteroides-Prevotella* group ($r_{Pearson} = -.18$). For further investigation we need to look into the link between maternal diet and breast milk composition.
Chapter 2

Introduction

2.1 Motivation

"When meditating over a disease, I never think of finding a remedy for it, but, instead, a means of preventing it.”
- Louis Pasteur

This thesis focuses on breastfeeding and gut bacteria - two fields with an important role in both disease prevention and promotion of well-being in global health (see 2.2.1, 2.5.1). With respect to different resource settings and support of the Sustainable Development Goals (UN 2015), I find something as universally applicable as breastfeeding particularly interesting.

With the profound impact of early childhood development in mind, I decided to investigate two physiological links between mother and infant through human milk:

- Diurnal Melatonin rhythm in human milk: A key focus question of the Baby Sleep Lab is the development of sleep in early childhood. A possible influence of the mother on their infant’s sleep is through hormones - such as melatonin - in breast milk.

- Association of maternal diet and BMI with infant gut bacteria: As infant’s gut microbiota undergoes rapid development within the first year, potentially influencing factors need to be identified and investigated. Maternal diet is an easily alterable factor and is therefore an accessible candidate for shaping the breastfed infant’s gut microbiota.
During the preparation of this thesis the following projects were proposed and outlined, but not executed:

- Developing an LC-MS analysis method able to track the diurnal variations of Cortisol, Cortisone and Melatonin in breast milk
- Investigating whether the human milk microbiome exhibits diurnal patterns

2.2 Human milk

2.2.1 Impact of breastfeeding

It has been shown that breastfeeding has a multitude of benefits for infants, mothers and societies. Victora et al. (2016) demonstrated with the Lives Saved Tool that “ [...] scaling up of breastfeeding to a near universal level could prevent 823’000 annual deaths in children younger than 5 years and 20’000 annual deaths from breast cancer.”. As a result, societies benefit from decreasing health care expenditures.

Reduced risk for infections of the lower respiratory tract in pre-school children with a breastfeeding period of 6 months or longer were shown by Tromp et al. (2017) based on the Dutch population-based prospective cohort "Generation R study". Further evidence of protective effects of breastfeeding against infectious diseases were found by Bowatte et al. (2015) against acute otitis media and by Lamberti et al. (2011) against diarrhoeal diseases. Human milk might also prevent common causes of infant deaths in high-income countries, e.g. from necrotizing enterocolitis (Sisk et al. 2007) and sudden infant death syndrome (Hauck et al. 2011; Vennemann et al. 2009). Breastfeeding has also been associated with improved performance in intelligence tests and with a positive effect on cognition in a randomised trial (Horta et al. 2015).

A recent study using data from the Canadian Healthy Infant Longitudinal Development (CHILD) cohort found that earlier weaning and supplementation with formula were associated with an increased risk of overweight at 12 months of age in a dose-dependent manner (Forbes et al. 2018; Mueller et al. 2018). The largest cluster-randomised controlled breastfeeding trial, the Promotion of Breastfeeding Intervention Trial (PROBIT), found no effect on adiposity at 6.5 years of age (Kramer et al. 2007).

Two main mechanisms have been proposed for the breast cancer protective effect (CGHFBC 2002) observed in mothers who have previously breastfed: differentiation of breast tissue and reduction in the lifetime number of ovulatory cycles. The human milk components alpha-lactalbumin and oleic acid can form the HAMLET (Human Alpha-lactalbumin Made LEthal to Tumor cells) complex which induces apoptosis in tumor cells while having no effect on differentiated cells (do Carmo França-Botelho 5...
et al. (2012). Duration of the breastfeeding period is the second most influential factor on birth interval length after contraceptive use. Becker et al. (2003) estimated the number of pregnancies avoided by lactational amenorrhea as high as 50% in countries with high breastfeeding rates such as Burkina Faso and Uganda. In a recent study from the Danish National Birth Cohort, Kirkegaard et al. (2018) showed an association between breastfeeding and lower maternal risk of cardiovascular diseases. Aune et al. (2014) suggest that there is a statistically significant inverse association between breastfeeding and maternal risk of type 2 diabetes. The meta-analysis of Luan et al. (2013) observed associations between breastfeeding and lower risks of epithelial ovarian cancer.

Opposite to widely held beliefs, research on the effect of breastfeeding on postpartum weight loss remains inconclusive (Neville et al. 2014; Sichieri et al. 2003). The Cochrane Database Systematic Review about "Breastfeeding or nipple stimulation for reducing postpartum haemorrhage in the third stage of labour" (Abedi et al. 2016) concluded that there is insufficient evidence to evaluate this effect. Potential underlying mechanisms are discussed in Saxton et al. 2014.

Horta et al. (2007) discussed general methodological challenges in studying the long-term effects of breastfeeding. Most evidence is derived from observational studies rather than randomized controlled trials (Lutter et al. 2012). Prominent exceptions were the PROBIT study and times where it was not considered unethically to randomly assign pre-term infants to formula or banked breast milk (Lucas et al. 1984). In the PROBIT study (Kramer et al. 2001) hospitals were randomly assigned to implement or not implement the Baby-Friendly Hospital Initiative. Breastfeeding duration and exclusivity were higher among those implementing the initiative. Regarding strength of evidence, prospective birth cohort studies should be considered the second best design after randomized controlled trials (Horta et al. 2007). Concato et al. 2000 showed that the results of meta-analysis from observational studies, with either cohort or case-control design, can be similar to that of randomized controlled trials. But the main issues in interpreting this kind of studies are self-selection, measurement errors, and residual confounding. The considered studies employed a broad range of assessment methods. For example, the assessment of breastfeeding (e.g. categorisation according to exclusive breastfeeding duration) and follow-up time differed substantially.
Table 2.1: Potential benefits of breastfeeding for infants, mothers and societies.

<table>
<thead>
<tr>
<th>Benefactor</th>
<th>Benefit</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infant</td>
<td>Infectious morbidity and mortality ↓</td>
<td>Victora et al. 2016</td>
</tr>
<tr>
<td></td>
<td>e.g. Respiratory tract infections ↓</td>
<td>Tromp et al. 2017</td>
</tr>
<tr>
<td></td>
<td>e.g. Acute otitis media ↓</td>
<td>Bowatte et al. 2015</td>
</tr>
<tr>
<td></td>
<td>e.g. Diarrhoea morbidity and mortality ↓</td>
<td>Lamberti et al. 2011</td>
</tr>
<tr>
<td></td>
<td>Dental malocclusions ↓</td>
<td>Peres et al. 2015</td>
</tr>
<tr>
<td></td>
<td>Performance in intelligence tests ↑</td>
<td>Horta et al. 2015</td>
</tr>
<tr>
<td></td>
<td>Overweight and obesity ↓</td>
<td>Forbes et al. 2018; Luoto et al. 2013; Mueller et al. 2018</td>
</tr>
<tr>
<td>Mother</td>
<td>Breast cancer ↓</td>
<td>CGHFBC 2002</td>
</tr>
<tr>
<td></td>
<td>Cardiovascular diseases ↓</td>
<td>Kirkegaard et al. 2018</td>
</tr>
<tr>
<td></td>
<td>Type 2 diabetes ↓</td>
<td>Aune et al. 2014</td>
</tr>
<tr>
<td></td>
<td>Ovarian cancer ↓</td>
<td>Luan et al. 2013</td>
</tr>
<tr>
<td>Society</td>
<td>Health care expenditure ↓</td>
<td>Rollins et al. 2016</td>
</tr>
</tbody>
</table>

2.2.2 Incidence of breastfeeding

Most mothers start to breastfeed at birth. Only three countries (France, Spain and the USA) with rates below 80% for ever breastfeeding were identified by Rollins et al. 2016. The report of UNICEF 2018 found that more than 1 in 5 babies in high-income countries are never breastfed, compared to 1 in 25 in low- and middle-income countries. Therefore, it is one of the few health positive behaviours more common in low-income compared to high-income countries. The incidence of breastfeeding in Switzerland is reported in the Swiss Infant Feeding Study (BLV 2016) by the Federal Food Safety and Veterinary Office. The study showed that 26% of children are exclusively breastfed within the first five to six months.

2.2.3 Anatomy and physiology of lactation

The exterior anatomy of the breast consists of the nipple (Papilla mammae), areola (Areola mammae) and Glands of Montgomery (Glandula areolaris). The interior anatomy of the non-lactating breast (Mamma non lactans) is dominated by adipose tissue. During pregnancy and lactation an increase in glandular tissue is induced, which results in a larger size of the lactating breast (Mamma lactans). The glandular tissue is divided into 15-20 gland lobules of alveoli, each lobule terminating into its own duct.
(Ductus lactifer). The lactiferous ducts are branched tubes up to 2 mm in diameter. Nipple and areola contain smooth musculature which contracts at stimulation, becomes erect and thus facilitates the feeding process (Geddes 2007; Spornitz 2010).

Figure 2.1: Overview of mammary gland physiology.

The lactation terminology used was defined by Neville et al. 2007; Pang et al. 2007.

- **Alveolar development** is associated with proliferative activity during the luteal phase of the reproductive cycle and in the early stages of pregnancy, which leads to formation of the milk secreting unit. The complexity of the terminal ductal lobular units increases through continued bud formation from the intralobular terminal ducts with recurrent menstrual cycles and pregnancy (Brisken 2002).

- **Secretory differentiation** occurs during pregnancy. Prolactin levels in the blood increase and stimulate the growth and development of the mammary tissue, e.g. the differentiation of mammary epithelial cells into milk producing cells (Brisken et al. 2010).

- **Secretory activation** is occurring around the time of parturition. After delivery, levels of progesterone and oestrogen fall rapidly, allowing milk secretion. Prolactin is controlling lactation in response to the suckling stimulus, among others, and shows a diurnal rhythm (Stern et al. 1990).

- **Milk ejection** is promoted by oxytocin which stimulates the mammary myoepithelial cells to contract, thereby promoting milk ejection during suckling (Sternlicht et al. 2006). This milk ejection reflex can occur coupled to the mother’s sensations and feelings.
• **Involution** of the mammary gland is characterized by apoptosis of the secretory epithelial cells undergoing programmed cell death and proteolytic degradation of the mammary gland basement membrane (Lund et al. 1996). Differences in involution between natural and forced weaning were observed (Silanikove 2014).

### 2.2.4 Composition of human milk

The composition of human milk changes within lactation period, day and feed.

Human milk changes over the lactation period from colostrum via transitional milk to mature milk. The colostrum contains more protein and less energy, fat and lactose than mature milk (Emmett et al. 1997; Gidrewicz et al. 2014). It is believed that colostrum has a mild laxative effect, promoting the passing of the infants’ meconium. Weaver et al. 1998 found no change with advancing lactation stage in the concentration of the main immunoglobulins in human milk IgA. Protein content decreases with postnatal age (Saarela et al. 2005). Fat (Khan et al. 2013) and energy content (Kociszewska-Najman et al. 2012) showed a diurnal pattern. Hampel et al. 2017 found variation in vitamin concentrations within feed and time of day.

Human Milk Oligosaccharides (HMOs) are the third most abundant solid component in human milk after lactose and lipids (Jeness 1979). More than 150 different HMOs have been identified. Infant formula oligosaccharides are structurally and functionally different from HMOs. Postulated beneficial effects include HMOs acting as prebiotics, antiadhesives, antimicrobials, and epithelial cell modulators (Bode 2012).

A potential factor influencing the composition is maternal diet. Data from Berger et al. 2018 suggests that maternal high-fructose corn syrup-sweetened beverage intake increased fructose concentrations in breast milk. Dunstan et al. 2004 observed that supplementation with fish oil, high in ω-3 polyunsaturated fatty acids, during pregnancy significantly alters early post-partum breast milk fatty acid composition. In addition to the macro- and micronutrients, human milk contains a microbiome (Section 2.2.5), maternal cells (e.g. leukocytes and stem cells (Briere et al. 2016)) and bioactive molecules (e.g. hormones and immunoglobulins).

### 2.2.5 Human milk microbiome

Previously, it was believed that human milk was sterile by default and only contaminated with bacteria from the mother’s skin or the infant’s mouth (West et al. 1979). Recently human milk was recognized as a source of commensal and potentially probiotic bacteria (Martín et al. 2003 2006). Cabrera-Rubio et al. 2012 found that aseptically collected colostrum already contains bacteria. Human milk shares bacteria
with the skin microbiome around the nipple (Grice et al. 2009; Hunt et al. 2011) and high levels of Streptococcus. Additionally, gut-associated obligate anaerobic bacteria and identical strains of bacteria in milk, blood and faecal samples of lactating women were identified by Jost et al. 2013; Perez et al. 2007 e.g. Bifidobacterium, Bacteroides, Parabacteroides and members of Clostridia. Two origins of the human milk bacteria have been proposed (Hinde et al. 2015): Skin contamination and retrograde flow (salivary backwash) as a result of the intra-oral vacuum dynamics of suckling (Geddes et al. 2008) could explain the high Streptococcus abundance, because Streptococcus is abundant in the infant saliva (Cephas et al. 2011). The more speculative translocation to milk through the bacterial entero-mammary pathway (Funkhouser et al. 2013; Rodríguez 2014) could explain the gut-associated obligate anaerobic bacteria.

The pathway of some bacteria crossing the intestinal epithelium and reaching the mammary gland and other locations is not yet understood. Proposed hypotheses involve dendritic cells and macrophages (Fernández et al. 2013). Dendritic cells can penetrate the gut epithelium to take up non-pathogenic bacteria from the gut lumen: They can open the tight junctions between intestinal epithelial cells, send dendrites outside the epithelium and sample bacteria. At the same time keeping the integrity of the epithelial barrier intact by expression of tight-junction proteins (Rescigno et al. 2001). Macpherson et al. 2004 showed that commensal bacteria can be retained for several days in mesenteric lymph nodes by intestinal dendritic cells. Antigen-stimulated cells can move from the intestinal mucosa to other mucosal surfaces, such as those of the respiratory tracts, salivary glands, and the lactating mammary gland (Murphy et al. 2017).

The human milk microbiome is one potential factor influencing the infant gut microbiota (Section 2.5.2). Pannaraj et al. 2017 have observed that infants receive about 25 % of their gut bacteria from breast milk and about 10 % from areolar skin during the first month of life. The human milk microbiome may contribute to reducing the incidence and severity of infections of the breastfed infant (Fernández et al. 2013) by various mechanisms, such as competitive exclusion (Olivares et al. 2006), production of antimicrobial compounds (Martín et al. 2006, Olivares et al. 2009), or improvement of the intestinal barrier function by increasing mucine production and reducing intestinal permeability (Olivares et al. 2006). For example, the incidence in gastrointestinal and upper respiratory tract infections was shown to be reduced by administration of Lactobacillus (Maldonado et al. 2012). Other potential impact by the human milk microbiome may lie within the metabolism of the infant. The glycobiome of some members of Lactobacillus and Bifidobacterium may help to nourish specific microbiota in the infant gut (Zivkovic et al. 2011). They may also contribute to infant digestion through the breakdown of sugars and proteins. Taking into account that the passage
of food through the gastrointestinal tract is shorter in infants than in adults and the pH value of the infantile stomach is higher, this effect would be particularly beneficial (Fernández et al. 2013). Strains of human milk *Lactobacillus* promote producing functional metabolites such as butyrate. Butyrate serves as main energy source for colonocytes and is modulating intestinal function, e.g. with an increase in faecal moisture, stool frequency, and volume (Gil-Campos et al. 2012; Maldonado et al. 2010).

Multiple factors are potentially influencing the human milk microbiome (Gomez-Gallego et al. 2016): mode of delivery, antibiotics, lactation stage, gestational age, health status, maternal diet and lifestyle, genetics, and geographical location. Cabrera-Rubio et al. 2012 observed higher microbial diversity and high prevalence of *Bifidobacterium* and *Lactobacillus* in breast milk following vaginal delivery compared to caesarian section. Maternal diet is a potential factor influencing the human milk microbiome as observed for fermented food (Albesharat et al. 2011) and fatty acid composition (Nishimura et al. 2014; Peng et al. 2009).

### 2.2.6 Human milk and sleep

Melatonin, tryptophan, nucleosides/nucleotides, and vitamin B12 are components of breast milk with sleep-promoting properties (Arslanoglu et al. 2012). In addition, Melatonin, tryptophan, and 5'-adenosine monophosphate and 5'-guanosine monophosphate nucleotides have been shown to exhibit diurnal patterns in human milk. However, associations between these diurnal patterns and the infants' sleep behaviour need to be researched. As full-term infants develop melatonin rhythmicity around three months of age (Attanasio et al. 1986; Kennaway et al. 1996; Vicente et al. 1989), human milk may act as a synchronizer after delivery.

### 2.3 Melatonin

#### 2.3.1 Impact of Melatonin

Melatonin is involved in various physiological processes including the sleep-wake cycle, seasonal adaptation, reproductive physiology and immunomodulation (Pandi-Perumal et al. 2008). The roles of melatonin seem to have evolved from its primary function as an antioxidant (Tan et al. 2010). The sleep promoting and circadian effects of melatonin are attributed to the melatonin receptors MT1 and MT2 present in the suprachiasmatic nucleus. It can reduce the sleep latency, even in small doses such as 0.1-0.3 mg/day (Pandi-Perumal et al. 2007). Besides Melatonin controlled release tablets for insomnia treatment in patients over 55 years of age (Hardeland 2012), other melatonergic drugs are approved by the European Medicines Agency. The non-selective
(MT1/MT2) melatonin receptor agonist and serotonergic 5-HT2c antagonist Agomelatine is the first reported melatonergic drug having anxiolytic and antidepressant effects (Pandi-Perumal et al. 2008). Diseases in which variations in production of endogenous melatonin production were shown are, among others, sleep disorders, Alzheimer’s disease, Parkinson’s disease, glaucoma and depression (Ekmekcioglu 2006; Pandi-Perumal et al. 2008).

### 2.3.2 Physiology of Melatonin

Melatonin (N-acetyl-5-methoxytryptamine) is produced by the pineal gland during the night. Specialized ganglion cells in the retina communicate variations in light to the suprachiasmatic nuclei of the hypothalamus. In turn, the suprachiasmatic nuclei induce down-regulation of melatonin production in the pineal gland (Claustrat et al. 2005).

Melatonin is synthesized (Figure 2.2) from the essential amino acid tryptophan to serotonin, which is converted into melatonin by a two-step process involving the two enzymes serotonin-N-acetyl transferase (NAT) and hydroxyindole-O-methyl transferase (HIOMT) (Klein et al. 1979).

The night peak defines biological night (Skene et al. 2006), shifts with age and is associated with large interindividual variance. Diurnal variations in expression of the clock gene PER1 depend on pineal gland and Melatonin secretion (Manev et al. 2006). There are multiple factors influencing Melatonin levels: light exposure, supplementation, stress (Kimata 2007), delivery mode and age. Pontes et al. 2007 observed that the increase in TNF-α after caesarean section resulted in suppression of the Melatonin night peak. Melatonin has also been identified in a large number of foods (Iriti et al. 2014; Meng et al. 2017).
2.4 Restaurant hypothesis

The "restaurant hypothesis" was first described by Conway et al. \cite{Conway2015b} Leatham-Jensen et al. \cite{Leatham-Jensen2012}. It is postulated that the gut microbiota is structured by nutrient availability into different nutrient-defined intestinal niches, so called "restaurants". The structure remains relatively invariant unless nutritional stress occurs (Nash et al. \cite{Nash2017}). The colonization resistance of \textit{Escherichia coli} was explained with this concept: \textit{E. coli} is present in "restaurants" with sugars of polysaccharide-degrading anaerobes (Conway et al. \cite{Conway2015a}). In order to survive and establish themselves, pathogens like
Salmonella must compete with the existing microbiota to obtain their own "restaurant" or will not be able to colonize over longer time frames (Lu et al. 2016).

Bokulich et al. 2016 found that the early gut microbiota is shaped by the availability of nitrogen and microbial strategies for obtaining nitrogen in the gut. What feeds the developing microbiota may determine its composition, and hence its function. Further implications for the "restaurant hypothesis" to human milk and influences of maternal diet to infant gut microbiota are discussed in Section 2.6.2.

2.5 Gut microbiota

2.5.1 Impact of the early gut microbiota

Abberations in the infant gut microbiota may stem from a broad range of influences and can cause an even broader range of issues: Caesarean sections raise the risk for pediatric obesity, antibiotic treatments may cause allergies to develop, and maternal obesity is associated with less microbial diversity and higher prevalence of pathogens (Nash et al. 2017). One of the most evident connections is the peculiar dependence of macrophage development on oligosaccharides produced by microbial metabolisms found by Luo et al. 2015. Epigenetic responses to the infant gut microbiota have also been observed (Berti et al. 2017). Whether or not the maternal phenotype has a meaningful impact is not clear.

The implications of these findings have been explored to find new pathways in the treatment of microbial abberations in the hope of preventing or lowering the chances of a variety of associated diseases (Marques et al. 2010). So far, the novel avenue of therapeutic manipulation of the gut microbiome involves faecal transplantation, which is unsuitable for infants. This limits the therapeutic methods to prebiotic and probiotic therapies, which show promise in reducing inflammation and associated diseases (van Best et al. 2015).

Based on data from the CHILD study, Arrieta et al. 2015 observed a decrease in Lachnospira, Veillonella, Faecalibacterium, and Rothia in infants at risk of asthma. Airway inflammation was developed after inoculation of germ-free mice with these four bacterial taxa. Azad et al. 2015 observed an association of low gut microbiota richness and an elevated Enterobacteriaceae/Bacteroidaceae ratio in early infancy with subsequent food sensitization.
2.5.2 Development of the early gut microbiota

Discussions about the initial colonization of the infant gut have been ongoing for almost 150 years (Küstner 1877). Since the second half of the last century, it was generally agreed upon that the foetus remained in a sterile state in utero (Escherich 1885). The advantages of molecular techniques led to the identification of rudimentary bacterial communities in the placenta, amniotic fluid and meconium from healthy pregnancies (Collado et al. 2016). These findings have sparked a new controversy between the “sterile womb paradigm” and the “in utero colonization hypothesis”.

The foetus is protected by anatomical, physical and immunological barriers. For example, toll-like receptors controlling molecular pattern recognition are present in the human placenta (Klaffenbach et al. 2005). Additionally, the female reproductive tract expresses antimicrobial peptides which serve as a chemical barrier to ascending infections (Yarbrough et al. 2015). The concentrations of some antimicrobial peptides are increased during late pregnancy, others are released into the amniotic fluid and to the offspring during delivery (Kai-Larsen et al. 2014). Bacterial pathogens such as *Listeria monocytogenes*, *Brucella abortus* and *Toxoplasma gondii* can overcome the barriers (Lecuit 2005) with severe consequences.

The main weaknesses of the “in utero colonization hypothesis” as summarized by Perez-Muñoz et al. 2017 are the insufficient detection limits used by the studies to detect “low-biomass” bacteria populations, lack of appropriate controls for contamination, and missing evidence for bacterial viability. The ability to deliver axenic animals via caesarean sections (Smith et al. 2007) supports the ”sterile womb hypothesis”.

Delivery remains a defining step in early microbiota development. As recent reviews have summarized (Kundu et al. 2017; Rodríguez et al. 2015), the first contact with contaminated maternal tissue determines the initial colonisation of the infant intestines. While the vaginal mode of delivery transmits bacteria typical of the vaginal microbiome (*Lactobacillus*, *Prevotella*), babies born by caesarian section display microbes more typical of skin (*Staphylococcus*, *Corynebacterium*, *Propionibacterium*) (Kundu et al. 2017). An established factor in the early development however is the gestational age of the newborn: Due to the sterile environment and higher incidence of antibiotic and general medical treatments, babies born preterm (earlier than 37th week) have a seriously impaired early colonisation, which leaves them more vulnerable to pathogens (Milani et al. 2017). Another influencing factor is the feeding mode (ibid.). The microbiota of breastfed infants becomes dominated by the obligate anaerobes *Bifidobacterium*, especially *B. bifidum*, *B. bifidum*, *B. breve*, and *B. longum subsp. infantis* (Turroni et al. 2012), influenced by the human milk oligosaccharides (Lewis et al. 2015). Additional influencing factors are antibiotic exposure (Nobel et al. 2015)
and environmental factors such as urbanisation (Ayeni et al. 2018). The family environment and genetic influences have been shown in twin studies (Stewart et al. 2005): genetically identical twins display significantly more similar microbiota than fraternal twins.

### 2.5.3 Diurnal rhythmicity of gut microbiota

Physiological function, disease expression and drug effects are associated with time of day, metabolome and oral microbiome as investigated by Skarke et al. 2017. To respond to daily variations in their environment, most life forms have evolved clock systems to adapt their physiology accordingly. Thaiss et al. 2015, 2016 observed that the microbiota, especially epithelial-adherent commensal bacteria, undergo diurnal oscillations in composition and function and that these oscillations are required for metabolic homeostasis of the host.

### 2.6 Diet and gut microbiota

#### 2.6.1 Diet and gut microbiota in adults

Diets are shaped by socio-economic status, cultural traditions, population growth, agriculture and many more. Potentially resulting shifts in the gut microbiome might have an impact on the host’s nutritional status and immune responses (Kau et al. 2011). A prominent example of the potential impact of nutrition and gut bacteria is *Zobellia galactanivorans* which can metabolise porphyran derived from marine red algae. Homologues of its porphyranase genes are present in the human gut bacterium *Bacteroides plebeius* and are strongly represented in the gut microbiota of the Japanese population. A possible explanation for this finding might be horizontal gene transfer from *Z. galactanivorans* to the human gut bacterium *B. plebeius* of a human host consuming contaminated food, and the host proceeding to transmit the altered bacteria to other Japanese (Hehemann et al. 2010).

The sensitivity of gut microbiota to changes in diet have been shown in animal models. Human faecal microbial community transplants in gnotobiotic mice supported the notion that the relative abundances of different gut bacteria species are sensitive to different foods (Goodman et al. 2011). Maukonen et al. 2015 provide an overview of observed associations between diet and their gut bacteria in adults. To further investigate the diet-microbiota link, Kau et al. 2011 concluded that dietary databases, new biomarkers of nutritional status, and quantitative phenotyping of the immune system are needed.
<table>
<thead>
<tr>
<th>Food item</th>
<th>Increase</th>
<th>Decrease</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fermented</td>
<td><em>Lactobacillus</em>, <em>Leuconostoc</em></td>
<td>F/B ratio, <em>Blautia</em></td>
<td>Han et al. 2015</td>
</tr>
<tr>
<td>Olive Oil</td>
<td><em>Bifidobacterium</em></td>
<td>-</td>
<td>Martín-Peláez et al. 2017</td>
</tr>
<tr>
<td>Seafood (PUFAs)</td>
<td><em>Lactobacillus</em></td>
<td><em>Coriobacteriaceae</em></td>
<td>Maukonen et al. 2015</td>
</tr>
<tr>
<td>Whole Grain</td>
<td><em>Bifidobacterium</em>, Butyrate producing</td>
<td>-</td>
<td>Rose 2014</td>
</tr>
<tr>
<td>Yogurt</td>
<td><em>Bifidobacterium</em></td>
<td><em>Bilophila wadsworthia</em></td>
<td>Burton et al. 2017</td>
</tr>
<tr>
<td>Low FODMAP</td>
<td>Mucus degrading</td>
<td><em>Bifidobacterium</em>,</td>
<td>Hill et al. 2017</td>
</tr>
<tr>
<td>Low carb (LW)</td>
<td><em>Oscillibacter</em></td>
<td><em>Eubacterium rectale</em>,</td>
<td>Walker et al. 2011</td>
</tr>
<tr>
<td>Vegetarian (SVD)</td>
<td><em>Bacteroides fragilis</em>, <em>Clostridia</em></td>
<td><em>Firmicutes</em>, <em>Enterobacteriaceae</em></td>
<td>do Rosario et al. 2016</td>
</tr>
</tbody>
</table>

### 2.6.2 Maternal diet and infant gut microbiota

While there are observations that an adult’s diet is shaping their gut microbiota, there is only limited research on correlations between a mother’s diet and her breastfed infant’s gut microbiota. Existing research demonstrated that a maternal high-fat diet, independent of obesity, is associated with an increase in *Enterococcus*, *Granulicatella* and *Lactococcus* and decrease in *Sutterella*, *Parabacteroides*, *Bacteroides* and *Comamonas* in the infant’s gut microbiome (Chu et al. 2016). Polyunsaturated fatty acids have been shown to restore the offspring’s gut microbiota in stressed animal models (Pusceddu et al. 2015). Macpherson et al. 2017 reviewed further consequences of maternal nutrition except for infant gut microbiota, e.g. decreased levels of maternal antibody that are transferred to the foetus via the placental neonatal Fc receptor due to malnutrition.

### 2.7 Body Mass Index

BMI categorisation is defined by WHO 2006 in which values from 18.5 to 24.9 kg m\(^{-2}\) constitute the normal weight range. Generally, the BMI methodology is limited by scaling issues, neglecting varying body parameters such as body fat, and a lack of penultimate
scientific evidence showing its connection to general health. A specific weakness of the BMI concept is that the mass increases to the third power of linear dimensions while the BMI is scaled down by the square of the height, thereby systematically disadvantaging taller individuals.

2.7.1 Body Mass Index and gut microbiota

Recently it was proposed to include gut microbiota into the two-way obesity model of the interplay between humans and their environment. Ongoing research aims to differentiate between obesogenic mechanisms of diet and those of microbiota.

Examples from animal models include studies with transplantation of the microbiota from murine weight loss models into germ-free mice which induced significant weight loss (Liou et al. 2013) and an increase in body fat without an increase in food consumption, suggesting that gut microbiota may play a role in metabolic efficiency (Bäckhed et al. 2004). Ley et al. 2006 showed that the Bacteroidetes phylum is underrepresented in proportion to Firmicutes in obese compared to lean people. Furthermore, observed changes in the intestinal microbiota and intestinal permeability may trigger inflammation in obese individuals (Cox et al. 2015). These findings were further supported by twin studies (Turnbaugh et al. 2009) with an additional increase in Actinobacteria in the obese group. Nadal et al. 2009 found a positive association between the Bacteroides-Prevotella group and weight loss.

Children born to obese mothers are at an increased risk for obesity themselves, but the mechanisms behind this association remain unclear (Paliy et al. 2014). A proposed hypothesis to link maternal and child obesity is the transmission of obesogenic bacteria from mother to child (Galley et al. 2014). Collado et al. 2010 found increased Bacteroides and Staphylococcus in infants’ faeces of overweight mothers during their first 6 months of life. Galley et al. 2014 also observed changes in the gut microbiota of infants born to obese mothers of higher socioeconomic status, e.g. in Parabacteroides, Oscillibacter, Eubacterium, Blautia.

2.8 Guidelines

2.8.1 Breastfeeding practices

WHO 2002 recommends exclusively breastfeeding infants for the first six months of life. WHO and UNICEF launched the Baby-friendly Hospital Initiative to promote implementation of the Ten Steps to Successful Breastfeeding strategy in facilities providing maternity and newborn services worldwide (WHO et al. 2018). This approach
encourages immediate and uninterrupted skin-to-skin contact, supports mothers to initiate breastfeeding as soon as possible after birth, discourages use of any food or fluids other than breast milk unless medically indicated. It also aims to enable rooming-in 24 hours a day in hospitals and allowing the mother to react to the feeding demands of their child rather than following a fixed schedule. The Cochrane Database Systematic Review (Kramer et al. 2012) stated that no apparent risks are associated with implementing exclusive breastfeeding for six months as a general health policy in both developing and developed-country settings.

### 2.8.2 Nutrition during the lactation period

The Swiss Federal Food Safety and Veterinary Office FSVO (Bundesamt für Lebensmittelsicherheit und Veterinärwesen BLV) published recommendations for nutrition during pregnancy and lactation (BLV 2017).

<table>
<thead>
<tr>
<th>Encouraged</th>
<th>Discouraged</th>
</tr>
</thead>
<tbody>
<tr>
<td>ω-3 fatty acids</td>
<td>Quinine</td>
</tr>
<tr>
<td>Iodine</td>
<td>Caffeine</td>
</tr>
<tr>
<td>Vitamin D</td>
<td>Alcohol</td>
</tr>
<tr>
<td>Unsweetened beverages</td>
<td>Venison</td>
</tr>
<tr>
<td>Balanced diet</td>
<td>Some fish species, e.g. marlin</td>
</tr>
</tbody>
</table>

Table 2.3: FSVO recommendations for nutrition during the lactation period.
Chapter 3

Aim

The following hypotheses were investigated:

1. Diurnal melatonin rhythm is observable in human milk.
   (a) A procedure for collection of human milk samples including study documents, materials and methods suitable for melatonin analysis can be developed and tested for feasibility in a pilot study.
   (b) ELISA analysis is a feasible method to analyse melatonin levels in human milk samples and observe its diurnal rhythm.

2. Differences in groups ordered by maternal food intake monitored with a Food Frequency Questionnaire are observable in the relative abundance of taxonomic units of interest in their breastfed infants’ gut bacteria.
   (a) Groups with a maternal diet high in whole grain show increased relative abundance of butyrate producing bacteria.
   (b) Groups with a maternal diet high in whole grain show increased relative abundance of *Bifidobacterium*.
   (c) Groups with a maternal diet high in yogurt show increased relative abundance of *Bifidobacterium*.
   (d) Groups with a maternal diet high in yogurt show decreased relative abundance of *Bilophila*.
   (e) Groups with a maternal diet high in olive oil show increased relative abundance of *Bifidobacterium*.
   (f) Groups with a maternal diet high in seafood show increased relative abundance of *Lactobacillus*. 
(g) Groups with a maternal diet high in seafood show decreased relative abundance of *Coriobacteriaceae*.

(h) Groups with a maternal diet high in fermented food show decreased relative abundance of *Blautia*.

(i) Groups with a maternal diet high in fermented food show increased relative abundance of *Lactobacillus*.

(j) Groups with a maternal diet high in fermented food show increased relative abundance of *Leuconostoc*.

(k) Groups with a maternal diet high in fermented food show decreased relative abundance of *Firmicutes/Bacteroidetes* ratio.

3. The maternal Body Mass Index shows an association with specific taxonomic units of their breastfed infants’ gut bacteria.

(a) Higher maternal BMI is associated with lower relative abundance of *Bacteroidetes* in infant gut bacteria.

(b) Higher maternal BMI is associated with higher relative abundance of *Firmicutes* in infant gut bacteria.

(c) Higher maternal BMI is associated with higher relative abundance of *Actinobacteria* in infant gut bacteria.

(d) Higher maternal BMI is associated with a higher *Firmicutes/Bacteroidetes* ratio in infant gut bacteria.

(e) Higher maternal BMI is associated with lower relative abundance of *Bacteroides-Prevotella* group in infant gut bacteria.

(f) Higher maternal BMI is associated with higher relative abundance of *Parabacteroides* in infant gut bacteria.

(g) Higher maternal BMI is associated with higher relative abundance of *Oscillibacter* in infant gut bacteria.

(h) Higher maternal BMI is associated with lower relative abundance of *Eubacterium* in infant gut bacteria.

(i) Higher maternal BMI is associated with lower relative abundance of *Blautia* in infant gut bacteria.
Chapter 4

Methods

4.1 Pilot Experiment

Figure 4.1: Procedure of the pilot study to observe diurnal melatonin rhythm in human milk.

4.1.1 Recruiting

The pilot participants were an exclusively breastfeeding mother and her 40 days, vaginally delivered old infant boy. The mother was currently not under medication.

4.1.2 Home visit

The mother received ten sampling kits, a breast pump (Medela Symphony), a 24h-breastfeeding diary, and an instruction sheet (see Appendix A). The sampling kits consisted of a labeled round-bottom polystyrene vial (Greiner Bio-One, 5 mL), a polyethylene plug (Greiner Bio-One), a single-use polyethylene pipette (Pastette®, 3 mL), and two aluminium foil sheets in individually labeled bags (Whirl-Pak®). The study staff gave an oral instruction on sample collection during the home visit.
4.1.3 Sampling

Hindmilk samples were collected at home with the provided breast pump at least every 4h over a 24h period, transferred into the vials, wrapped into aluminium foil, and stored at -20 °C (household freezer) to avoid UV and heat degradation. Sample time and all feeding times were recorded in the 24h-breastfeeding diary.

4.2 Human milk analysis

The analysis was conducted at the Genetic Diversity Centre Zurich.

4.2.1 Sample preparation

Before analysis, the samples were thawed to room temperature with help of a vortexer (BR-2000 Vortexer Bio-Rad) and centrifuged (Thermo Scientific Heraeus™ Megafuge™ 40 R) at 3'300 g for 15 min at 4 °C. Subsequently, the heavier aqueous phase was placed on conditioned extraction columns and extracted with methanol (Honeywell Riedel-de Haen, for HPLC ≥ 99 %). The methanol was evaporated in a evaporator centrifuge (Eppendorf Concentrator 5301) at 30 °C for three hours in multiple steps.

4.2.2 ELISA analysis of Melatonin

The samples were analysed in duplicates with a commercial enzyme-linked immunosorbent assay (ELISA) kit (IBL International, RE54021) for human serum and plasma samples. This particular kit was previously used for milk analysis by Asher et al. 2015; Kimata 2007; Kollmann et al. 2008. The pipetting scheme shown in Figure 4.2 was used. Six standards including one negative control (CAL A-F), two controls (CTL 1-2) and eight samples (01-08) were prepared. In the assay, biotinylated Melatonin competed with the sample Melatonin at the binding sites of anti-rabbit IgG. After incubation for 18 hours, the amount of antibody bound biotinylated Melatonin was determined with the enzyme conjugate streptavidin alkaline phosphatase which converts the substrate p-nitrophenyl phosphate to the yellow p-nitrophenol. The absorbance was measured at 405 nm (Tecan Spark®).
4.3 Infant gut bacteria study

4.3.1 Participants

129 mother-infant pairs participated in this study. 9 participants had to be excluded due to low bacteria counts (n=7), ablactation during measurement (n=1), and uncertain sample identity (n=1). Therefore, 120 samples of infants at 3 months of age (range 76-109 days, mean 95.4 ± 6.5 days) and their mothers (range 22-43, mean 33.7 ± 3.8 years) were considered.

The participants of the study were recruited between July 2016 and May 2018. Direct recruitment was done at the postnatal care of the University Hospital Zurich. Additionally, recruitment was conducted through address lists of families with infants provided by municipalities and distribution of flyers. To ensure an informed consent process, participants recruited during postnatal care were called for further explanation of the study procedure when their children reached two months of age.

Inclusion criteria for the study were vaginal delivery, predominantly breastfeeding at the first examination (at three months of age) and good general health status. Exclusion criteria were birth before the 37th or after the 42nd pregnancy week, birth weight of less than 2'500 g, antibiotic or medication use influencing sleep/wake rhythms, family background of narcolepsy, psychosis or bipolar disorders, prenatal contact with recreational drugs, prenatal infections or known or suspected abuse of drugs or alcohol by the mother. Examinations were planned in a way to avoid time zone changes over the prior month and vaccination in the last two weeks prior.
4.3.2 Process

Interested participants filled out two screening questionnaires and one consent form online before the first measurements. The participants were visited at their homes before the first measurement to give detailed instructions along with the provided material. Measurements lasted 10 days in total. Preferably during the last two days of the measurement, two stool samples from two different diapers were collected, planned to be within 48 hours (range 0.5-77.4 hours, mean 24.8 ± 15.8 hours) before the collection of the samples by study staff. A sampling kit was prepared with plastic gloves, pipettes, spatulae, gauze, coded tubes and bags. Parents noted the time of the sample collection, the previous bowel movement and unusual occurrences. The samples were stored in a household refrigerator until collection. A member of the study team visited the family after each measurement to pick up the material and to discuss any anomalies during data acquisition. The transport from participants to University Hospital was done in a cooling box at 0-5 °C. The sample was weighed and aliquots of minimum 200 mg were made. These samples were stored at -50 °C until analysis. Additionally, the maternal diet was monitored with a self-reported online Food Frequency Questionnaire (see Appendix A).

4.4 Longitudinal study and further involvement

This thesis is based on the longitudinal cohort study Schlafmuster und Darmbakterien von gesunden Säuglingen und Kleinkindern (SDEGU, ”Sleep Behavior and Gut Microbiota in Healthy Infants”), where infants were examined at three, six and twelve months of age. In addition to the analysed parameters, infant sleep was measured via actigraphy and documented in a sleep diary (10 days per timepoint). For a subgroup of the participants, maternal and/or paternal sleep patterns were also recorded. A high-density sleep electroencephalography (hdEEG) was measured of 22 participants at 6 months of age for two hours at the participants’ homes. The study was approved by the cantonal ethics commission Zurich.
4.5 Analysis of infant gut bacteria

4.5.1 16S rRNA gene analysis

About 200 mg of the faecal samples were extracted using PowerSoil® DNA Isolation Kit (MOBIO Laboratories, Carlsbad, CA, USA) with the following modifications from the manufacturer: Prior to DNA extraction, samples were placed into the PowerBead tubes and heat treated at 65°C for 10 min and then at 95°C for 10 min. Subsequently, solution C1 was added and bead-beating performed in FastPrep (MP Biomedicals, Santa Ana, CA, USA) using 3 cycles of 15 s each, at a speed of 6.5 m/s.

16S rRNA gene-amplicon sequencing of the V3 region (Ovreas et al. 1997) was performed using primers designed with adapters for the Nextera Index Kit® (Illumina, CA, USA) NXt_338_F and NXt_518_R. The first PCR for amplification profiling, second PCR for barcoding, amplicon library purification, and sequencing were conducted according to Krych et al. 2018. FastQC (v0.11.2, Andrews 2010) was used to quality check the generated data of two times 151 bp (pair-ended). seqtk (Shen et al. 2016) and FLASH (v1.2.11, Magoc et al. 2011) were used with the following param-
eters for trimming and merging of overlapping forward and reverse reads respectively: minimum overlap of 15, maximum overlap of 300 and maximum mismatch density of 0.25. Primer regions allowing for an error rate of 0.01 were identified and trimmed with Cutadapt (v1.12, Martin 2011). Reads with a mean quality of less than 20 or containing ambiguous nucleotides were removed during quality filtering with PrinSeq (Schmieder et al. 2011). Operational Taxonomic Units (OTUs) at 97 % sequence identity were defined from the generated and de-noised amplicons with usearch (v10.0.240, UNoise3, Edgar 2010). The 16S rRNA gene analysis databases SINTAX (Edgar 2016) and SILVA (v128) were used for taxonomic predictions.

4.5.2 Data analysis

The obtained taxonomic gut bacteria data was analysed with the phyloseq R library (v1.24.0, McMurdie et al. 2013). After reading in the data, filtering and taxonomic visualisations, statistical analysis (Section 4.6) was performed.

4.6 Statistical analysis

Two mothers had to be excluded due to missing BMI information. Therefore, 118 mother-infant pairs were considered for calculations involving BMI.

The statistical analysis was conducted with R (v3.5.0, RStudio version 1.1.453) according to the workflow visualised in Figure 4.4. After visualisation of the data in histograms and scatterplots, it was examined for anomalies in skewness and group populations. Depending on the measurement scale of the independent variable (interval scale for BMI data, ordinal scale for FFQ data) linear regression analysis or ANOVA workflow was conducted.

The effect size of significant regression pairs was determined according to Cohen 1988 with the Pearson’s correlation coefficient $r$.

Depending on the success of the Shapiro-Wilk test for normal distribution and the Levene’s test for homogeneity of variance, the ANOVA, Welch one-way test or Kruskal-Wallis test were employed.
Figure 4.4: Outline of the statistical analysis workflow.
Chapter 5

Results

5.1 Pilot Experiment

5.1.1 Sampling

The pilot participant collected eight breast milk samples over a 24h period (see Figure 5.1). Feedback about the study design was received during the home visit for sample collection. She reported that sampling was more convenient at the end rather than in the middle of a feed. Furthermore, she preferred to couple the sampling times to the infant’s feeding behaviour instead of planning in advance to a strict 4h rhythm. Nevertheless, she adhered sufficiently closely to the projected schedule. She proposed to decrease the sample volume in order to avoid the use of a breast pump.

5.1.2 ELISA analysis of Melatonin

As seen in Figure 5.1 the melatonin concentration curve peaks at 11 pg/mL between 2 and 5 am. During daytime, the melatonin level in the milk samples dropped below the ELISA detection limit (1.6 pg/mL). Between minimum and maximum, the measured concentrations rose monotonously. Together with the intermediate values, the diurnal rhythm that was to be expected (Claustrat et al. 2005; Kimata 2007) was observed.
5.2 Maternal diet, Body Mass Index and infant gut bacteria

5.2.1 Infant gut bacteria at 3 months

The phylogenetic tree (Figure 5.2) shows that the most abundant phyla were *Firmicutes, Actinobacteria, Bacteroidetes* and *Proteobacteria*, which were detected in substantial abundances in the overwhelming majority of samples.
Figure 5.2: Phylogenetic tree of OTUs with abundance >500 on genus level. The length of the bars denotes the number of samples which contained detectable amounts of bacteria belonging to the respective genus.
Figure 5.3: Abundance histograms of the investigated bacteria clusters on phylum level. The data takes the form of heavily right-skewed Poisson distributions, as is typical for count-type data obtained from natural abundances.

First analysis of abundance distribution with histograms (Figure 5.3 and 5.4) led to exclusion of taxonomic units with more than 50% of data points below detection limit or skewness $> 5$.

Therefore the phyla *Firmicutes, Actinobacteria* and *Bacteroidetes*, the class *Clostridia*, the family *Coriobacteriaceae* and the genera *Bacteroides, Bifidobacterium, Parabacteroides* and *Staphylococcus* were investigated.
5.2.2 Food Frequency Questionnaire distributions

The histograms of the six food items of interest are shown in Figure 5.5. Distributions with high skewness (> 1), kurtosis (> 3) and unpopulated intake categories were excluded from further analyses.
5.2.3 Body Mass Index distribution

From the self-reported height and weight of the mothers, their BMI was calculated. The distribution of different BMI scores is shown in Figure 5.6. It takes the form of a significantly right-skewed bell curve (mean = 23.5 kg/m$^2$, sd = 3.7 kg/m$^2$, skewness = 1.1).
5.2.4 Association of maternal diet and infant gut bacteria

Since the assumption of normality does not hold in any of the investigated pairs (Shapiro-Wilk test all $p < .05$) and could not be restored using square root or logarithmic transformations, it is necessary to relax the testing model. Therefore, the Kruskal-Wallis test was used to investigate significant differences between the self-reported food intake groups. There were no significant differences between the FFQ groups for all investigated clusters using the Kruskal-Wallis test (all $p > .05$).

5.2.5 Regression of infant gut bacteria and maternal BMI

The regression of maternal BMI and the phylum *Actinobacteria* of the infant gut bacteria is significant ($F(116)=4.0, p = .05, r = .18$). The Pearson correlation coefficient $r = .18$ shows a positive correlation between maternal BMI and *Actinobacteria* the infantile gut bacteria with a small effect size according to Cohen [1988]. $R^2$ of .033 shows that maternal BMI can account for only 3.3 % of the variation in the relative abundance of the infantile *Actinobacteria* phylum. Maternal BMI and the relative abundance of the *Bacteroides-Prevotella* were also significantly correlated with a small effect size ($F(116)=4.4, p = .039, r = -.19$).
The correlation between maternal BMI and *Bacteroidetes* \( F(116)=3.9, p = .05, r = -.18 \) and with *Firmicutes*, *Firmicutes/Bacteroidetes* ratio and *Parabacteroides* (all \( p > .05 \)) were not significant.

Figure 5.7: Linear regression analysis of the correlation between maternal BMI and the infantile gut bacteria phyla *Actinobacteria* and *Bacteriodetes*, and the *Bacteroides-Prevotella* group. The regression models for *Firmicutes*, *Firmicutes/Bacteroidetes* ratio and *Parabacteroides* were not significant with \( p > .05 \).
Chapter 6

Discussion

6.1 Diurnal Melatonin rhythm in human milk

ELISA analysis and the developed procedure are feasible methods to monitor the diurnal Melatonin rhythm. Therefore, hypotheses 1a and 1b are fulfilled.

Different influencing factors on the melatonin level have to be considered in designing a study procedure and interpreting the data. The most prominent factor is probably light exposure, especially regarding electronics use at late hours and the associated blue light exposure.

The pilot participant proposed to decrease the sample volume in order to avoid the use of a breast pump. From my experience, the reduction in sample volume will most likely interfere with the centrifugation step for phase separation. I propose that a minimum of 2 mL is still needed for a clean phase separation, resulting in a sufficient amount of aqueous phase for ELISA testing (> 0.5 mL). As long as the mother can reliably express a minimum of 2 mL of breast milk via hand expression, I see no reason to insist on using a breast pump.

The melatonin night peak in human milk suggests an importance of night milk and chrononutrition, i.e. giving expressed milk to the infant at a similar time of day as collection. To account for Melatonin resorption and retention in infants, it is preferable to test the Melatonin in infant saliva in addition to the milk levels.
6.2 Association of maternal diet and BMI with infant gut bacteria

Based on the results of the Kruskal-Wallis tests, there was no significant difference between the FFQ groups for any of the investigated bacteria. Therefore, the hypotheses 2a-k must be rejected.

A possible factor for the poor prediction power may be the Food Frequency Questionnaire, since we did not control for any personal biases or other confounding variables. Moreover, the focus on rather specific bacteria groups may have contributed. Instead of focusing on specific taxonomic units, a more general approach like Principal coordinates analysis (PCoA) or diversity analysis may be more suited for this rather unexplored field of study. While we did exclude delivery mode, antibiotic exposure and feeding mode, it is known that a large number of other factors influence gut bacteria. Therefore, this current design suffers from the multiple comparison problem: Due to the large number of influences on the observed variable (bacteria abundance), it is extraordinarily difficult to infer information from singular axes. Caesarian section was an exclusion criteria in this particular study, but would be an interesting factor to explore and compare against vaginal deliveries. I suggest to test the hypothesis that thorough breastfeeding can be beneficial to mitigate the adverse microbial effects of delivery by caesarian section.

The maternal BMI showed an association with Actinobacteria and the Bacteroides-Prevotella group. Therefore, hypotheses 3c and 3e are fulfilled albeit with a small effect size, the remaining hypotheses must be rejected.

As phyla are high taxonomic units, no predictions about functionality are possible. This severely limits our ability to assign certain outcomes to gut microbiota composition. An example for the profound impact that improper scoping can have on the results is the fact I was able to correlate the maternal BMI with the Bacteroides-Prevotella group but not with its phylum Bacteroidetes. As I did not generally filter for a maximum relative abundance of the operational taxonomic units, the investigated groups might contain faultily characterised bacteria.

6.3 Conclusions

The diurnal Melatonin rhythm in human milk is one example of breast milk as a link between mother and infant. Since ELISA has proven to be a simple and sufficiently precise method to track the diurnal Melatonin rhythm, this study design is a feasible option for future investigations of the importance of Melatonin for infant sleep behaviour.
Infant gut bacteria seem to be much more resilient against variations to maternal diet than I thought. Whether or not this is an intrinsic property of itself or breast milk needs further investigation. To conduct a more reliable investigation, more potentially influencing factors must be identified and controlled for.

### 6.4 Outlook

The most obvious link to examine next is the influence of human milk Melatonin on infants’ sleep behaviour. Of particular interest might be the effect of breastfeeding behaviour (e.g. number and duration of night feedings) and the absolute breast milk Melatonin levels (e.g. potentially supplementation). This avenue is probably straightforward since we already have experience with monitoring sleep behaviour via actigraphy and sleep diaries. For the next pilot experiment, hand expression should be sufficient and more convenient for sampling for most mothers.

To investigate the potential impact of the alterable factor maternal diet further, I propose to analyse breast milk for changes by maternal diet first before jumping from mother to infant. In addition to bacteria, other human milk components should be considered in this effort as well. Particular components of interest due to their metabolic implications are among others polyunsaturated fatty acids or fructose.

The enormous progress of this field over the past decade has paved the way to elucidate the underlying mechanisms and possible impacts of early childhood nutrition. Right now, many of the new insights are still inconclusive. However, by making steady progress in dissecting each and every one of the possible links between mother and infant, we can hopefully raise support for breastfeeding and help future mothers make informed decisions founded in hard science for their own and their infants’ health.
Chapter 7

Acknowledgements

I would like to thank

- Dr. Salome Kurth,
- Sarah Schoch, and
- PD Dr. Vivianne Otto for their constant support and advice.
- My parents, Lukas, Mara and Juli for always believing in me.
References

Abedi, P et al. (2016). “Breastfeeding or nipple stimulation for reducing postpartum haemorrhage in the third stage of labour”. In: Cochrane Database of Systematic Reviews. DOI: 10.1002/14651858.CD010845.pub2.


Becker, S et al. (2003). “Estimation of births averted due to breastfeeding and increases in levels of contraception needed to substitute for breastfeeding”. In: Journal of Biosocial Science 35, pp. 559–574. DOI: 10.1017/S0021932003005996.


Burton, K et al. (2017). “Probiotic yogurt and acidified milk similarly reduce postprandial inflammation and both alter the gut microbiota of healthy, young men”. In: British Journal of Nutrition 117, pp. 1312–1322. DOI: 10.1017/S0007114517000885.

Cephas, K et al. (2011). “Comparative Analysis of Salivary Bacterial Microbiome Diversity in Edentulous Infants and Their Mothers or Primary Care Givers Using Pyrosequencing”. In: *PLOS ONE* 6. DOI: [10.1371/journal.pone.0023503](https://doi.org/10.1371/journal.pone.0023503).


Collado, M et al. (2016). “Human gut colonisation may be initiated in utero by distinct microbial communities in the placenta and amniotic fluid”. In: *Scientific Reports* 6. DOI: [10.1038/srep23129](https://doi.org/10.1038/srep23129).


Funkhouser, L et al. (2013). ‘Mom Knows Best: The Universality of Maternal Microbial Transmission’. In: PLOS Biology 11. DOI: 10.1371/journal.pbio.1001631.


Gil-Campos, M et al. (2012). “Lactobacillus fermentum CECT 5716 is safe and well tolerated in infants of 1–6 months of age: A Randomized Controlled Trial”. In: Pharmacological Research 65, pp. 231–238. DOI: 10.1016/j.phrs.2011.11.016.


Hinde, K et al. (2015). “Mother’s littlest helpers”. In: Science 348, pp. 1427–1428. DOI: 10.1126/science.aac7436.


45

Hunt, K et al. (2011). “Characterization of the Diversity and Temporal Stability of Bacterial Communities in Human Milk”. In: PLOS ONE 6. DOI: 10.1371/journal.pone.0021313

Iriti, M et al. (2014). “Melatonin in Mediterranean diet, a new perspective”. In: Journal of the Science of Food and Agriculture 95. DOI: 10.1002/jsfa.7051


Khan, S et al. (2013). “Variation in Fat, Lactose, and Protein Composition in Breast Milk over 24 Hours”. In: Journal of Human Lactation 29, pp. 81–89. DOI: 10.1177/0890334412448841


Kirkegaard, H et al. (2018). “Breastfeeding and later maternal risk of hypertension and cardiovascular disease - The role of overall and abdominal obesity”. In: Preventive Medicine 114, pp. 140–148. DOI: 10.1016/j.ypmed.2018.06.014


Liou, A et al. (2013). “Conserved Shifts in the Gut Microbiota Due to Gastric Bypass Reduce Host Weight and Adiposity”. In: Science Translational Medicine 5. DOI: 10.1126/scitranslmed.3005687.
Lu, C et al. (2016). “Commensal E. coli induced colonization resistance against mucosal Salmonella infection”. In: J Immunol 196. URL: http://www.jimmunol.org/content/196/1_Supplement/66.11.short.


Martín, R et al. (2006). “*Lactobacillus salivarius* CECT 5713, a potential probiotic strain isolated from infant feces and breast milk of a mother–child pair”. In: *International Journal of Food Microbiology* 112, pp. 35–43. DOI: [10.1016/j.ijfoodmicro.2006.06.011](https://doi.org/10.1016/j.ijfoodmicro.2006.06.011)


Meng, X et al. (2017). “Dietary Sources and Bioactivities of Melatonin”. In: *Nutrients* 9. DOI: [10.3390/nu9040367](https://doi.org/10.3390/nu9040367)


Mueller, E et al. (2018). “Breast milk, formula, the microbiome and overweight”. In: *Nature Reviews Endocrinology*. DOI: [10.1038/s41574-018-0066-5](https://doi.org/10.1038/s41574-018-0066-5)


Nobel, Y et al. (2015). “Metabolic and metagenomic outcomes from early-life pulsed antibiotic treatment”. In: Nature Communications 6. DOI: [10.1038/ncomms8486](https://doi.org/10.1038/ncomms8486)


Pandi-Perumal, S et al. (2007). “Drug Insight: the use of melatonergic agonists for the treatment of insomnia—focus on ramelteon”. In: Nature Clinical Practice Neurology 3, pp. 221–228. DOI: [10.1038/ncpneuro0467](https://doi.org/10.1038/ncpneuro0467)


Rescigno, M et al. (2001). “Dendritic cells express tight junction proteins and penetrate gut epithelial monolayers to sample bacteria”. In: *Nature Immunology* 2, pp. 361–367. DOI: 10.1038/86373.
Silanikove, N (2014). “Natural and abrupt involution of the mammary gland affects differently the metabolic and health consequences of weaning”. In: Life Sciences 102, pp. 10–15. DOI: 10.1016/j.lfs.2014.02.034

Sisk, P et al. (2007). “Early human milk feeding is associated with a lower risk of necrotizing enterocolitis in very low birth weight infants”. In: Journal of Perinatology 27, pp. 428–433. DOI: 10.1038/sj.jp.7211758

Skarke, C et al. (2017). “A Pilot Characterization of the Human Chronobiome”. In: Scientific Reports 7. DOI: 10.1038/s41598-017-17362-6


Smith, K et al. (2007). “Use of axenic animals in studying the adaptation of mammals to their commensal intestinal microbiota”. In: Seminars in Immunology 19, pp. 59–69. DOI: 10.1016/j.smim.2006.10.002


Stewart, J et al. (2005). “Investigations into the influence of host genetics on the predominant eubacteria in the faecal microflora of children”. In: Journal of Medical Microbiology 54, pp. 1239–1242. DOI: 10.1099/jmm.0.46189-0.


Thaiss, C et al. (2016). “Microbiota Diurnal Rhythmicity Programs Host Transcriptome Oscillations”. In: Cell 167, pp. 1495–1510. DOI: 10.1016/j.cell.2016.11.003


Turroni, F et al. (2012). “Diversity of Bifidobacteria within the Infant Gut Microbiota”. In: PLOS ONE 7. DOI: 10.1371/journal.pone.0036957


Weaver, L et al. (1998). “Human milk IgA concentrations during the first year of lactation”. In: Archives of Disease in Childhood 78, pp. 235–239. DOI: 10.1136/adc.78.3.235


Appendix A

Study documents

A.1 24h-breastfeeding diary
A.2 Instruction sheet
A.3 Food Frequency Questionnaire
## 24-Stunden-Still-Protokoll

<table>
<thead>
<tr>
<th>00:00</th>
<th>01:00</th>
<th>02:00</th>
<th>03:00</th>
<th>04:00</th>
<th>05:00</th>
<th>06:00</th>
<th>07:00</th>
<th>08:00</th>
<th>09:00</th>
<th>10:00</th>
<th>11:00</th>
<th>12:00</th>
<th>13:00</th>
<th>14:00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:00</td>
<td>16:00</td>
<td>17:00</td>
<td>18:00</td>
<td>19:00</td>
<td>20:00</td>
<td>21:00</td>
<td>22:00</td>
<td>23:00</td>
<td>00:00</td>
<td>01:00</td>
<td>02:00</td>
<td>03:00</td>
<td>04:00</td>
<td>05:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>06:00</td>
<td>07:00</td>
<td>08:00</td>
<td>09:00</td>
<td>10:00</td>
<td>11:00</td>
<td>12:00</td>
<td>13:00</td>
<td>14:00</td>
<td>15:00</td>
<td>16:00</td>
<td>17:00</td>
<td>18:00</td>
<td>19:00</td>
<td>20:00</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Probenentnahme (Uhrzeit):

Sample 01 ____ : ____
Sample 02 ____ : ____
Sample 03 ____ : ____
Sample 04 ____ : ____
Sample 05 ____ : ____
Sample 06 ____ : ____
Sample 07 ____ : ____
Sample 08 ____ : ____
Sample 09 ____ : ____
Sample 10 ____ : ____

Bitte tragen Sie alle Stillzeiten wie folgt ein:

<table>
<thead>
<tr>
<th>00:00</th>
<th>01:00</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sample xy ____ : ____

Anleitung Muttermilch-Probe

Bitte entnehmen Sie zwischen dem 26.06.18 und 27.06.18 innerhalb von 24 Stunden mindestens sechs Muttermilch-Proben folgendermassen:

1. **Zeitpunkt**

2. **Abpumpen**
   In der Mitte eines Stillvorganges entweder mit Milchpumpe oder per Hand abpumpen.

3. **Transfer**
   5 mL Muttermilch mit Pipette in entsprechendes Probenröhrchen transferieren. Bitte beachten Sie die korrekte Probennummer.

4. **Umhüllen**
   Probenröhrchen mit Alu-Folie umhüllen.

5. **Lagerung**
   Das umhüllte Probenröhrchen zurück in den entsprechenden Plastiksack geben und schnellstmöglich im Gefrierfach (ca. -20 °C) lagern.

6. **Protokoll**
   Probenentnahme im 24-Stunden-Still-Protokoll vermerken.

---

**Anleitung 24-Stunden-Still-Protokoll**

Bitte vermerken Sie alle Stillzeiten im Protokoll.

---

**Kontakt:**
Viktoria Gastens, MSc Student
viktoria.gastens@usz.ch
Tel. 076 441 70 42

Baby Schlaflabor / CRPP Sleep and Health
Department of Pulmonology
University Hospital of Zurich
Rämistr. 100
CH-8091 Zürich

Tel. 043 253 01 31
www.bsl.uzh.ch
Nehmen Sie mehr als 75% der täglichen Kalorien der Nahrung durch Ernährungs-Shakes zu sich?

- Ja
- Nein
- Manchmal

Während einer normalen Woche, wie häufig konsumieren Sie ...

<table>
<thead>
<tr>
<th>Fleisch/Eier?</th>
<th>Nie</th>
<th>Ab und zu (1-2 Mal)</th>
<th>Oft (3-5 Mal)</th>
<th>Täglich</th>
</tr>
</thead>
<tbody>
<tr>
<td>selbstgekochte Nahrung (ohne Fertiggerichte)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fertiggerichte (z.B. Tiefkühl-Pizza oder Nudelgerichte)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nahrung aus einem Restaurant, einschliesslich Take-out?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mindestens 2 Portionen Vollkornprodukte an einem Tag? (1 Produkt = 1 Scheibe 100% Vollkornbrot, 1 Tasse Vollkornmuesli, 3-4 Vollkorncrackers, 1/2 Tasse brauner Reis oder Vollweizen Pasta)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mindestens 2-3 Portionen Früchte an einem Tag? (1 Portion = 1/2 Tasse Früchte; 1 mittelgrosse Frucht, 1 dl 100% Fruchtsaft)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mindestens 2-3 Portionen Gemüse (inkl. Kartoffeln) an einem Tag? (1 Portion = 1/2 Tasse Gemüse, Kartoffeln, 1 Tasse rohes blattartiges Gemüse)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fermentiertes Gemüse/Pflanzen pro Tag (Milchsäuregärung)? (1 Portion = 1/2 Tasse Sauerkraut, Kimchi oder fermentiertiertes Gemüse oder 1 Tasse Kombucha)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>kultivierte Milchprodukte, gesäuert Milchprodukte oder Joghurt am Tag? (1 Portion = 1 Tasse Joghurt)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
mindestens 2 Portionen Milch oder Käse am Tag? (1 Portion = 1 Tasse Milch oder Joghurt; 50 g Käse)
Milch-Ersatz? (Sojamilch, Laktosefreie Milch, Mandelmilch, etc.)
gefrorene Desserts (Glace/Soft Ice/Milkshakes, Sorbet, Gefrorenes Joghurt, etc.)?

<table>
<thead>
<tr>
<th></th>
<th>Nie</th>
<th>1x pro Woche</th>
<th>1-2 Mal</th>
<th>3-5 Mal</th>
<th>Täglich</th>
</tr>
</thead>
<tbody>
<tr>
<td>rotes Fleisch?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rotes Fleisch mit grössem Fettanteil (z.B. Entrecote, T-Bone Steak, Hamburger, Rippchen, Speck, etc.)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Geflügel (Poulet oder Truthahn)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meeresfrüchte (Fisch, Hummer, Krabbe, Shrimp, etc.)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>gesalzene Snacks (Kartoffelchips, Nacho Chips, Mais Chips, Popcorn mit Butter, Pommes Frites, etc.)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>zuckerhaltige Süsiggkeiten (Kuchen, Guetzli, Süßgebäck, Donuts, Muffins, Schokolade) mindestens einmal pro Tag?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mit Olivenöl Gekochtes?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ganze Eier (keine Produkte mit nur Eiweiss)?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5dl oder mehr von zuckerhaltigen Süßgetranken oder Frucht-Drinks/Punsch (Coca Cola, Sprite, Fanta, Sirup etc. jedoch nicht 100% Fruchtsaft)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mindestens 1 Liter Wasser an einem Tag?</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Wieviele verschiedenen Arten Früchte und Gemüse essen Sie in einer typischen Woche? (Eine Suppe mit Erbsen, Karotten und Kartoffeln gilt als 3)

- Weniger als 5
- 6 – 10
- 11 – 20
- 21 – 30
- Mehr als 30

Bitte schreiben Sie hier alles andere auf, wovon Sie denken, dass es möglicherweise Ihre persönlichen Körperekken beeinflusst:
Appendix B

R Code
require(phyloseq)
require(ggplot2)
require(vegan)
library(dplyr)

# Read in data
otufile <- "p551_180314_180406_16S_OTU_Count_Sintax.txt"
mapfile <- "p551_180314_180406_16S_MapFile.txt"
treefile <- "p551_180314_180406_16S_OTU_CLU.tre"
refseqfile <- "p551_180314_180406_16S_OTU.fa"

# 97% OTU clustering data prep
dOTU <- import_qiime(otufilename = otufile,
  mapfilename = mapfile,
  treefilename = treefile,
  refseqfilename = refseqfile)
dOTU

# Sample counts
sort(sample_sums(dOTU))

# remove samples with very low counts
dOTU <- subset_samples(dOTU,
  X.SampleID != "SO81-03_16S" &
  X.SampleID != "none_16S" &
  X.SampleID != "SO63-03_16S" &
  X.SampleID != "SO72-03_16S")
dOTU <- subset_samples(dOTU,
  X.SampleID != "SO69-06_16S" &
  X.SampleID != "SO66-06_16S" &
  X.SampleID != "SI17-03_16S" &
  X.SampleID != "SO32-12_16S")
dOTU <- subset_samples(dOTU,
  X.SampleID != "SO18-12_16S" &
  X.SampleID != "SO96-03_16S" &
  X.SampleID != "SO87-03_16S" &
  X.SampleID != "SI30-03_16S")

# remove samples with double run
dOTU <- subset_samples(dOTU,
  X.SampleID != "SO15-06_16S" &
  X.SampleID != "SO36-12_16S" &
  X.SampleID != "SO34-06_16S" &
  X.SampleID != "SO58-06_16S" &
  X.SampleID != "SO100-06_16S" &
  X.SampleID != "SO04-12_16S")

# remove samples with double run
dOTU <- subset_samples(dOTU,
  X.SampleID != "SO01-06a_16S" &
  X.SampleID != "SO04-06_16S" &
  X.SampleID != "SO04-12_16S" &
  X.SampleID != "SO15-06_16S" &
  X.SampleID != "SO36-12_16S" &
  X.SampleID != "SO34-06_16S" &
  X.SampleID != "SO58-06_16S" &
  X.SampleID != "SO100-06_16S" &
  X.SampleID != "SO04-12_16S")
X.SampleID != "S009-03a_16S" &
X.SampleID != "S010-03a_16S" &
X.SampleID != "S011-03a_16S" &
X.SampleID != "S012-03b_16S"

# remove ablactated participants
dOTU <- subset_samples(dOTU, X.SampleID != "S022-03_16S")

# remove samples with missing BMI information --> nur bei BMI Rechnungen
# dOTU <- subset_samples(dOTU, X.SampleID != "S068-03_16S" & X.SampleID != "S089-03_16S")

# remove sample with uncertain identity
dOTU <- subset_samples(dOTU, X.SampleID != "S129-03_16S")

# subset data (Subset negative samples)
dOTU.S <- subset_samples(dOTU, Group == "S")
dOTU.N <- subset_samples(dOTU, Group == "N")

# subset only 3 Months
dOTU.S.TP03 <- subset_samples(dOTU, TimePoint == "TP03")

# relative abundance filtering 3 months
dOTU.S.TP03.ra <- transform_sample_counts(dOTU.S.TP03,
   function(x) { 100 * (x / sum(x)) })

otu_table(dOTU.S.TP03.ra)

# zero (reset) relative abundances below 0.01
otu_table(dOTU.S.TP03.ra)[otu_table(dOTU.S.TP03.ra) < 1] <- 0

# remove OTUs with only zero relative abundance
dOTU.S.TP03.ra <- prune_taxa(taxa_sums(dOTU.S.TP03.ra) > 0, dOTU.S.TP03.ra)
notu_table(dOTU.S.TP03.ra) <- floor(otu_table(dOTU.S.TP03.ra))
dOTU.S.TP03.ra # 123 taxa

# filter OTUs that appear less than 3 times (per group/group replicates) 3 months
f3 <- genefilter_sample(dOTU.S.TP03.ra, filterfun_sample(function(x) x > 3))
d.f3 <- prune_taxa(f3, dOTU.S.TP03.ra)

ntaxa(dOTU.S.TP03.ra) - ntaxa(d.f3) # number of OTUs filtered
d.f3 # 63 taxa
# GMB Data, May 2018, Phyloseq
source("gmb_phyloseq.R")

# Include only infants of 3 months of age
gmb_3m <- dOTU.S.TP03

# Bifidobacterium
gmb_3m_bifido <- subset_taxa(gmb_3m, Genus == "Bifidobacterium")
bifido <- psmelt(gmb_3m_bifido)
bifido_sum <- aggregate(Abundance ~ Genus + ID, data=bifido, sum)

# Lactobacillus
gmb_3m_lacto <- subset_taxa(gmb_3m, Genus == "Lactobacillus")
lacto <- psmelt(gmb_3m_lacto)
lacto_sum <- aggregate(Abundance ~ Genus + ID, data=lacto, sum)

# Leuconostoc
gmb_3m_leuco <- subset_taxa(gmb_3m, Genus == "Leuconostoc")
leuco <- psmelt(gmb_3m_leuco)
leuco_sum <- aggregate(Abundance ~ Genus + ID, data=leuco, sum)

# Bilophila
gmb_3m_bilo <- subset_taxa(gmb_3m, Genus == "Bilophila")
bilo <- psmelt(gmb_3m_bilo)
bilo_sum <- aggregate(Abundance ~ Genus + ID, data=bilo, sum)

# Coriobacteriaceae
gmb_3m_corio <- subset_taxa(gmb_3m, Family == "Coriobacteriaceae")
corio <- psmelt(gmb_3m_corio)
corio_sum <- aggregate(Abundance ~ Family + ID, data=corio, sum)

# Blautia
gmb_3m_blau <- subset_taxa(gmb_3m, Genus == "Blautia")
blau <- psmelt(gmb_3m_blau)
blau_sum <- aggregate(Abundance ~ Genus + ID, data=blau, sum)

# Parabacteroides
gmb_3m_para <- subset_taxa(gmb_3m, Genus == "Parabacteroides")
para <- psmelt(gmb_3m_para)
para_sum <- aggregate(Abundance ~ Genus + ID, data=para, sum)

# Oscillibacter
gmb_3m_osci <- subset_taxa(gmb_3m, Genus == "Oscillibacter")
osci <- psmelt(gmb_3m_osci)
osci_sum <- aggregate(Abundance ~ Genus + ID, data=osci, sum)

# Eubacterium
gmb_3m_eubact <- subset_taxa(gmb_3m, Genus == "Eubacterium")
eubact <- psmelt(gmb_3m_eubact)
eubact_sum <- aggregate(Abundance ~ Genus + ID, data=eubact, sum)

# Firmicutes
gmb_3m_firm <- subset_taxa(gmb_3m, Phylum == "Firmicutes")
firm <- psmelt(gmb_3m_firm)
firm_sum <- aggregate(Abundance ~ Phylum + ID, data=firm, sum)

# Bacteroidetes
gmb_3m_bact <- subset_taxa(gmb_3m, Phylum == "Bacteroidetes")
bact <- psmelt(gmb_3m_bact)
bact_sum <- aggregate(Abundance ~ Phylum + ID, data=bact, sum)

# Proteobacteria
gmb_3m_proteo <- subset_taxa(gmb_3m, Phylum == "Proteobacteria")
proteo <- psmelt(gmb_3m_proteo)
proteo_sum <- aggregate(Abundance ~ Phylum + ID, data=proteo, sum)

# Firmicutes : Bacteroidetes ratio
fbratio <- data.frame(levels(firm$ID), rep(0, length(levels(firm$ID))))
colnames(fbratio) <- c("ID", "FBratio")
for (i in levels(firm$ID)) {
  fbratio[fbratio$ID == i, ]$FBratio <- firm_sum[firm_sum$ID==i, ]$Abundance /
                                bact_sum[bact_sum$ID==i, ]$Abundance
}

# Akkermansia
gmb_3m_acker <- subset_taxa(gmb_3m, Genus == "Akkermansia")
acker <- psmelt(gmb_3m_acker)
acker_sum <- aggregate(Abundance ~ Genus + ID, data=acker, sum)

# Clostridia
gmb_3m_clos <- subset_taxa(gmb_3m, Class == "Clostridia")
clos <- psmelt(gmb_3m_clos)
clos_sum <- aggregate(Abundance ~ Class + ID, data=clos, sum)

# Staphylococcus
gmb_3m_staph <- subset_taxa(gmb_3m, Genus == "Staphylococcus")
staph <- psmelt(gmb_3m_staph)
staph_sum <- aggregate(Abundance ~ Genus + ID, data=staph, sum)

# Roseburia
gmb_3m_rose <- subset_taxa(gmb_3m, Genus == "Roseburia")
rose <- psmelt(gmb_3m_rose)
rose_sum <- aggregate(Abundance ~ Genus + ID, data=rose, sum)

# Bacteroides
gmb_3m_bactg <- subset_taxa(gmb_3m, Genus == "Bacteroides")
bactg <- psmelt(gmb_3m_bactg)
bactg_sum <- aggregate(Abundance ~ Genus + ID, data=bactg, sum)

# Prevotella
gmb_3m_prevo <- subset_taxa(gmb_3m, Genus == "Prevotella")
prevo <- psmelt(gmb_3m_prevo)
prevo_sum <- aggregate(Abundance ~ Genus + ID, data=prevo, sum)
# Bacteroides-Prevotella group

```r
gmb_3m_bactprevo <- subset_taxa(gmb_3m, 
    Genus == "Bacteroides" | Genus=="Prevotella")
bactprevo <- psmelt(gmb_3m_bactprevo)
bactprevo_sum <- aggregate(Abundance ~ ID, data=bactprevo, sum)
```

# butyrate producing bacteria:
# Roseburia, Eubacterium, Clostridia

```r
gmb_3m_butyrate <- subset_taxa(gmb_3m, 
    Genus=="Roseburia" | Genus=="Eubacterium" | 
    Class=="Clostridia")
butyrate <- psmelt(gmb_3m_butyrate)
butyrate_sum <- aggregate(Abundance ~ Phylum + ID, data=butyrate, sum)
```

# Actinobacteria in %

```r
gmb_3m_actino <- subset_taxa(gmb_3m, Phylum == "Actinobacteria")
actino <- psmelt(gmb_3m_actino)
actino_sum <- aggregate(Abundance ~ Phylum + ID, data=actino, sum)
```

# Questionnaire T1, 07/2018

```r
qt1 <- read.csv("data_sdegu_2018-05-14_12-10.csv", header = TRUE)
qt1 <- filter(qt1, 
    SERIAL != "S022" & SERIAL != "S028" & SERIAL != "S063" & 
    SERIAL != "S072" & SERIAL != "S081" & SERIAL != "S087" & 
    SERIAL != "S096" & SERIAL != "S117" & SERIAL != "S129" & 
    SERIAL != "S130")
qt1 <- filter(qt1, SERIAL=="S068" & SERIAL=="S089")
```

# FFQ - Food Frequency Questionnaire

```r
ffq <- qt1[, grep("SERIAL|FM32_*|FM25_01|FM26_01", colnames(qt1))]
# FFQ reduced to food items of interest
# Vollkorn 05, Fermentiertes 08, Milchprodukte 09
# Käse 10, Meeresfische 16, Olivenöl 19
ffq_6fooditems <- ffq[, grep("SERIAL|FM32_((05)|(08)|(09)|(10)|(16)|(19))", 
    colnames(ffq))]
```

# Weight and Size

```r
bmi <- ffq[, grep("SERIAL|FM25_01|FM26_01", colnames(ffq))]
bmi[, 5] <- rep("0", length(bmi[, 4]))
colnames(bmi) <- c(colnames(bmi)[1:5], "BMI", "BMI_25")
bmi$BMI_level[bmi$BMI < 18.5] <- "<18.5"
bmi$BMI_level[bmi$BMI >= 18.5 & bmi$BMI < 25] <- "18.5-24.9"
bmi$BMI_level[bmi$BMI >= 25 & bmi$BMI < 30] <- "25.0-29.9"
bmi$BMI_level[bmi$BMI >= 30] <- ">30"
```

```r
glom <- tax_glom(gmb_3m, taxrank="Phylum")
```
dat <- psmelt(glom)
dat$Phylum <- as.character(dat$Phylum)
Phylum_abundance <- aggregate(Abundance~ID + Phylum, dat, FUN=sum)
Phylum_abundance <- cast(Phylum_abundance, ID+Phylum)

# Add a column with summary of all Phyla
Phylum_abundance$Sums <- rowSums(Phylum_abundance)

# Relative Abundance of Actinobacteria, Firmicutes, Bacteroidetes, Proteobacteria
percentage <- mutate(Phylum_abundance, Actinobacteria = Actinobacteria / Sums)
percentage <- mutate(percentage, Firmicutes = Firmicutes / Sums)
percentage <- mutate(percentage, Bacteroidetes = Bacteroidetes / Sums)
percentage <- mutate(percentage, Proteobacteria = Proteobacteria / Sums)

# loeschen
percentage$Cyanobacteria <- NULL
percentage$Deinococcus-Thermus <- NULL
percentage$Fusobacteria <- NULL
percentage$Saccharibacteria <- NULL
percentage$Tenericutes <- NULL
percentage$Verrucomicrobia <- NULL

# BMI vs Bacteroidetes
bmi_bact <- merge(x = bact_sum,
y = bmi[, c("SERIAL", "BMI_level")],
by.x = "ID",
by.y = "SERIAL")

# BMI vs Actinobacteria
bmi_actino <- merge(x = actino_sum,
y = bmi[, c("SERIAL", "BMI_level")],
by.x = "ID",
by.y = "SERIAL")

# Whole grain vs butyrate producing bacteria
butyrate_wholegrain <- merge(x = butyrate_sum,
y = ffq_6fooditems[, c("SERIAL", "FM32_05")],
by.x = "ID",
by.y = "SERIAL")

# Whole Grain vs Bifidobacteria
bifido_wholegrain <- merge(x = bifido_sum,
y = ffq_6fooditems[, c("SERIAL", "FM32_05")],
by.x = "ID",
by.y = "SERIAL")

# Yogurt vs Bifidobacteria
bifido_yogurt <- merge(x = bifido_sum,
y = ffq_6fooditems[, c("SERIAL", "FM32_09")],
by.x = "ID",
by.y = "SERIAL")
```r
# (c) 2018 Viktoria Gastens

library(ggplot2)
library(cowplot)
library(moments)

source("read_data.R")

# Whole grain
h1 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_05)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Whole Grain",
  subtitle="skew = -0.1, kurtosis = 2.2") +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
  labels=c("Never", "< 1", "1-2", "3-5", "7"))

# Fermented
h2 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_08)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Fermented",
  subtitle="skew = 0.5, kurtosis = 3.8") +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
  labels=c("Never", "< 1", "1-2", "3-5", "7"),
  limits=c(NA, 5))

# Milk
h3 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_09)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Milk",
  subtitle="skew = -0.6, kurtosis = 2.5") +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
  labels=c("Never", "< 1", "1-2", "3-5", "7"))

# Cheese
h4 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_10)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Cheese",
  subtitle="skew = -1.2, kurtosis = 3.7") +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
  labels=c("Never", "< 1", "1-2", "3-5", "7"))

# Seafood
h5 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_16)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Seafood",
  subtitle="skewness = -0.1, kurtosis = 2.1") +
```
# Olive oil

```r
h6 <- ggplot(data=ffq_6fooditems, aes(ffq_6fooditems$FM32_19)) +
  geom_histogram(binwidth=1) +
  ylim(0, 80) +
  labs(x="Meals per week", y="Count", title="Olive Oil",
       subtitle="skew = -0.6, kurtosis = 2.7") +
  scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
                   labels=c("Never", "< 1", "1-2", "3-5", "7"))

p <- plot_grid(h1, h2, h3, h4, h5, h6, ncol=3)

cowplot::ggsave("ffq_histogram.png", width=12, height=6)
```

---

### bmi_histogram.R

```r
library(ggplot2)
library(cowplot)
library(moments)
source("read_data.R")

bmi_plot <- ggplot(data=bmi, aes(bmi$BMI, fill=bmi$BMI_level)) +
  geom_histogram(binwidth=0.5, center=0.25) +
  xlim(16, 40) +
  scale_y_continuous(breaks=seq(0, 12, 2)) +
  labs(x="Body Mass Index (kg/m2)", y="Number of Mothers") +
  scale_fill_manual(values=c("blue", "green", "orange", "red"),
                     name="",
                     labels=c("underweight", "normal weight", "pre-obesity", "obesity")) +
  geom_vline(xintercept=18.5, linetype="dashed") +
  geom_vline(xintercept=25, linetype="dashed") +
  geom_vline(xintercept=30, linetype="dashed")

ggsave("bmi_histogram.png")
```

---

### bacteria_histogram.R

```r
library(ggplot2)
library(cowplot)
library(moments)
source("read_data.R")

# Phyla histogram
```
# Actinobacteria

p1 <- ggplot(data=actino_sum, aes(actino_sum$Abundance)) +
  geom_histogram(boundary=0) +
  xlim(0, 400000) +
  labs(x="Abundance", y="Count", title="Actinobacteria",
       subtitle="n(0) = 0, skew = 1.5")

# Firmicutes

p2 <- ggplot(data=firm_sum, aes(firm_sum$Abundance)) +
  geom_histogram(boundary=0) +
  xlim(0, 400000) +
  labs(x="Abundance", y="Count", title="Firmicutes",
       subtitle="n(0) = 0, skew = 3.1")

# Bacteroidetes

p3 <- ggplot(data = bact_sum, aes(bact_sum$Abundance)) +
  geom_histogram(boundary=0) +
  xlim(0, 400000) +
  labs(x="Abundance", y="Count", title="Bacteroidetes",
       subtitle="n(0) = 0, skew = 2.9")

# Proteobacteria

p4 <- ggplot(data = proteo_sum, aes(proteo_sum$Abundance)) +
  geom_histogram(boundary=0) +
  xlim(0, 400000) +
  labs(x="Abundance", y="Count", title="Proteobacteria",
       subtitle="n(0) = 0, skew = 1.4")

p <- plot_grid(p1, p2, p3, p4, ncol=2)
cowplot::ggsave("bacteria_histogram_phylum.png", width=10, height=8)

# Class histogram

c1 <- ggplot(data=clos_sum, aes(clos_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 50000, 100000, 150000), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Clostridium",
       subtitle="n(0) = 0, skew = 4.2")

c <- plot_grid(c1, ncol=2)
cowplot::ggsave("bacteria_histogram_class.png", width=10, height=5)

# Genus histogram

# Akkermansia

g1 <- ggplot(data=akker_sum, aes(akker_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 10000, 20000, 30000), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Akkermansia",
       subtitle="n(0) = 92, skew = 8.4")

# Bacteroides

g13 <- ggplot(data=bactg_sum, aes(bactg_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 100000, 200000, 300000, 400000), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Bacteroides",
       subtitle="n(0) = 0, skew = 8.4")
limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Bacteroides",
      subtitle="n(0) = 2, skew = 3.4")

# Bifidobacterium
g2 <- ggplot(data=bifo_sum, aes(bifo_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 100000, 200000, 300000, 400000),
                    limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Bifidobacterium",
       subtitle="n(0) = 0, skew = 1.5")

# Bifidobacterium

g3 <- ggplot(data=bilo_sum, aes(bilo_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 200, 400, 600, 800), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Bilophila",
       subtitle="n(0) = 103, skew = 6.4")

# Blautia

g4 <- ggplot(data=blau_sum, aes(blau_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 20000, 40000, 60000), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Blautia",
       subtitle="n(0) = 62, skew = 10.3")

# Eubacterium

g5 <- ggplot(data=eubact_sum, aes(eubact_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 1000, 2000, 3000, 4000),
                    limits=c(0.5, NA)) +
  scale_y_continuous(breaks=c(0, 1, 2)) +
  labs(x="Abundance", y="Count", title="Eubacterium",
       subtitle="n(0) = 114, skew = 10.0")

# Lactobacillus

g6 <- ggplot(data=lacto_sum, aes(lacto_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 20000, 40000, 60000), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Lactobacillus",
       subtitle="n(0) = 45, skew = 10.6")

# Leuconostoc

g7 <- ggplot(data=leuco_sum, aes(leuco_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 5, 10), limits=c(0.5, 10)) +
  labs(x="Abundance", y="Count", title="Leuconostoc",
       subtitle="n(0) = 116, skew = 5.2")

# Oscillibacter

g8 <- ggplot(data=osci_sum, aes(osci_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 200, 400, 600), limits=c(0.5, NA)) +
# Parabacteroides

g9 <- ggplot(data=para_sum, aes(para_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 50000, 100000, 150000),
                     limits=c(0.5, NA),
                     labels=c("1", "50000", "100000", "150000")) +
  labs(x="Abundance", y="Count", title="Parabacteroides",
       subtitle="n(0) = 36, skew = 6.1")

# Prevotella

g10 <- ggplot(data=prevo_sum, aes(prevo_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 20, 40, 60), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Prevotella",
       subtitle="n(0) = 78, skew = 5.0")

# Roseburia

g11 <- ggplot(data=rose_sum, aes(rose_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 2000, 4000, 6000, 8000),
                     limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Roseburia",
       subtitle="n(0) = 86, skew = 10.7")

# Staphylococcus

g12 <- ggplot(data=staph_sum, aes(staph_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 50, 100, 150, 200), limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Staphylococcus",
       subtitle="n(0) = 23, skew = 4.4")

# Family histogram

# Coriobacteriaceae
f1 <- ggplot(data=corio_sum, aes(corio_sum$Abundance)) +
  geom_histogram(boundary=0.5) +
  scale_x_continuous(breaks=c(1, 5000, 10000, 150000, 20000),
                     limits=c(0.5, NA)) +
  labs(x="Abundance", y="Count", title="Coriobacteriaceae",
       subtitle="n(0) = 19, skew = 4.3")

p <- plot_grid(c1, g1, g13, g2, g3, g4, g5, g6, g7, g8, g9, g10, g11, g12, f1, ncol=3)
cowplot::ggsave("bacteria_histogram_genus.png", width=20, height=15)

ffq_bacteria_boxplots.R

# (c) 2018 Viktoria Gastens

library(car)
library(ggplot2)
library(pastecs)
library(psych)

source("read_data.R")

bmilevels <- c("underweight", "normalweight", "preadipositas", "adipositas")

# Whole Grain vs Bifidobacteria
bifido_wholegrain.plot <- ggplot(data=bifido_wholegrain,
                                  aes(x=FM32_05, y=Abundance, group =FM32_05, fill=bifido_wholegrain$FM32_05)) +
                        labs(x="Meals per week", y="Abundance", title="Whole grain vs Bifidobacteria") +
                        scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
                                           labels=c("Never", "< 1", "1-2", "3-5", "7")) +
                        geom_boxplot() +
                        guides(fill=FALSE)

# Yogurt vs Bifidobacteria
bifido_yogurt.plot <- ggplot(data=bifido_yogurt,
                              aes(x=FM32_09,y=Abundance, group=FM32_09, fill=bifido_yogurt$FM32_09)) +
                       labs(x="Meals per week",y="Abundance", title="Yogurt vs Bifidobacteria") +
                       scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
                                          labels=c("Never", "< 1", "1-2", "3-5", "7")) +
                       geom_boxplot() +
                       guides(fill=FALSE)

ggsave("bifido_yogurt.png", plot=bifido_yogurt.plot)

# Olive Oil vs Bifidobacteria
bifido_oil.plot <- ggplot(data=bifido_oil,
                          aes(x = FM32_19, y=Abundance, group=FM32_19, fill=bifido_oil$FM32_19)) +
                     labs(x="Meals per week",y="Abundance", title="Olive Oil vs Bifidobacteria") +
                     scale_x_continuous(breaks=c(1, 2, 3, 4, 5),
                                        labels=c("Never", "< 1", "1-2", "3-5", "7")) +
                     geom_boxplot() +
                     guides(fill=FALSE)

ffq_kruskalwallis.R

# (c) 2018 Viktoria Gastens
library(car)

source("read_data.R")

# Adding bifido vs whole grain
percentage <- merge(x=percentage, y=bifido_wholegrain, by.x="ID", by.y="ID")
percentage$Genus <- NULL
percentage$FM32_05 <- NULL

names(percentage)[names(percentage) == "Abundance"] <- "Bifido"
percentage <- mutate(percentage, Bifido = Bifido / Sums)

# Adding bifido vs yogurt
percentage <- merge(x=percentage, y=bifido_yogurt, by.x="ID", by.y="ID")
percentage$Genus <- NULL
percentage$Abundance <- NULL

# Adding bifido vs olive oil
percentage <- merge(x=percentage, y=bifido_oil, by.x="ID", by.y="ID")
percentage$Genus <- NULL
percentage$Abundance <- NULL

# Adding butyrate producing vs whole grain
percentage <- merge(x=percentage, y=butyrate_wholegrain, by.x="ID", by.y="ID")
percentage$Genus <- NULL
names(percentage)[names(percentage) == "Abundance"] <- "Butyrate"
percentage <- mutate(percentage, Butyrate=Butyrate / Sums)

# Shapiro-Wilk test: normal distribution of residuals
# P > 0.05: SUCCEEDS
# P < 0.05: FAILS
bifido_wholegrain.shapiro <- shapiro.test(residuals(object=aov(Bifido ~ FM32_05, data = percentage)))
bifido_yogurt.shapiro <- shapiro.test(residuals(object=aov(Bifido ~ FM32_09, data = percentage)))
bifido_oil.shapiro <- shapiro.test(residuals(object=aov(Bifido ~ FM32_19, data = percentage)))
butyrate_wholegrain.shapiro <- shapiro.test(residuals(object=aov(Butyrate ~ FM32_05, data = percentage)))

# Kruskal Test
# P > 0.05: SUCCEEDS
# P < 0.05: FAILS
bifido_wholegrain.kruskal <- kruskal.test(Bifido ~ as.factor(FM32_05), data = percentage)
bifido_yogurt.kruskal <- kruskal.test(Bifido ~ as.factor(FM32_09), data = percentage)
bifido_oil.kruskal <- kruskal.test(Bifido ~ as.factor(FM32_19), data = percentage)
butyrate_wholegrain.kruskal <- kruskal.test(Butyrate ~ as.factor(FM32_19), data = percentage)

bmi_regression.R

# (c) 2018 Viktoria Gastens
```r
library(dplyr)
library(ggplot2)
library(cowplot)

# Adding BMI
percentage <- merge(x=percentage, y=bmi, by.x="ID", by.y="SERIAL")
percentage$FM25_01 <- NULL
percentage$FM26_01 <- NULL

# Adding F/B ratio
percentage <- merge(x=percentage, y=fbratio, by.x="ID", by.y="ID")

# Adding Bacteroidetes Prevotella Group
percentage <- merge(x=percentage, y=bactprevo_sum, by.x="ID", by.y="ID")
colnames(percentage) <- c(colnames(percentage)[1:length(colnames(percentage)) - 1],
"Bact_Prevo")
percentage <- mutate(percentage, Bact_Prevo = Bact_Prevo / Sums)
percentage$Genus.x <- NULL

# Adding Parabacteroides
percentage <- merge(x=percentage, y=para_sum, by.x="ID", by.y="ID")
colnames(percentage) <- c(colnames(percentage)[1:length(colnames(percentage)) - 1],
"Para")
percentage <- mutate(percentage, Para = Para / Sums)
percentage$Genus.x <- NULL

# Logistic regressions
bmi_actino.lm <- lm(Actinobacteria ~ BMI, data=percentage)
bmi_actino.lmplot <- ggplot(data=percentage, aes(x=Actinobacteria, y=BMI)) +
  geom_point() +
  geom_smooth(method="lm") +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="Actinobacteria")

bmi_bact.lm <- lm(Bacteroidetes ~ BMI, data=percentage)
bmi_bact.lmplot <- ggplot(data=percentage, aes(x=Bacteroidetes, y=BMI)) +
  geom_point() +
  geom_smooth(method="lm") +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="Bacteroidetes")

bmi_firm.lm <- lm(Firmicutes ~ BMI, data=percentage)
bmi_firm.lmplot <- ggplot(data=percentage, aes(x=Firmicutes, y=BMI)) +
  geom_point() +
  geom_smooth(method="lm") +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="Firmicutes")

bmi_fbratio.lm <- lm(FBratio ~ BMI, data=percentage)
bmi_fbratio.lmplot <- ggplot(data=percentage, aes(x=FBratio, y=BMI)) +
  geom_point() +
  geom_smooth(method="lm") +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="FBRatio")
```

bmi_bactprevo.lm <- lm(Bact_Prevo ~ BMI, data=percentage)
bmi_bactprevo.lmplot <- ggplot(data=percentage, aes(x=Bact_Prevo, y=BMI)) +
  geom_point() +
  geom_smooth(method="lm") +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="Bacteroides-Prevotella")

bmi_para.lm <- lm(Para ~ BMI, data=percentage)
bmi_para.lmplot <- ggplot(data=percentage, aes(x=Para, y=BMI)) +
  geom_point() +
  labs(x="Relative Abundance",
       y=expression(paste("BMI [kg / m"^2,"]")),
       title="Parabacteroides")

# Save plot
bmi.lmplot <- plot_grid(bmi_actino.lmplot, bmi_bact.lmplot, bmi_bactprevo.lmplot,
                        bmi_firm.lmplot, bmi_fbratio.lmplot, bmi_para.lmplot,
                        ncol=3)
cowplot::ggsave("bmi_bacteria.png", width=15, height=7)

bmi_pearson.R

# (c) 2018 Viktoria Gastens
source("bmi_regression.R")

# Pearson correlation tests
bmi_actino.pearson <- cor.test(x = percentage$BMI,
                                y = percentage$Actinobacteria,
                                method = "pearson")
bmi_bact.pearson <- cor.test(x = percentage$BMI,
                              y = percentage$Bacteroidetes,
                              method = "pearson")
bmi_firm.pearson <- cor.test(x = percentage$BMI,
                              y = percentage$Firmicutes,
                              method = "pearson")
bmi_fbratio.pearson <- cor.test(x = percentage$BMI,
                                 y = percentage$FBratio,
                                 method = "pearson")
bmi_bactprevo.pearson <- cor.test(x = percentage$BMI,
                                      y = percentage$Bact_Prevo,
                                      method = "pearson")
bmi_para.pearson <- cor.test(x = percentage$BMI,
                              y = percentage$Para,
                              method = "pearson")
milk_stats.R

# (c) 2018 Viktoria Gastens

library(ggplot2)
library(drc)

# Read data
melatonin <- read.csv("ELISA_melatonin_2018-06-29.csv", header = TRUE)

melatonin.cal <- melatonin[grepl("CAL*", melatonin$Sample),]
melatonin.ctlsam <- melatonin[grepl("CTL*|P004*", melatonin$Sample),]

# log conc
melatonin.cal$logconc <- log10(melatonin.cal$Melatonin_pg_mL)

## calibration curve
# IBL: cubic-spline-method, 4-parameter-Analyse lin-log, logit-log Berechnung
# von plot zu ggplot
calibration.plot <- ggplot(data=melatonin.cal,
aes(melatonin.cal$logconc, melatonin.cal$OD)) +
  geom_point() +
  labs(x="log Melatonin [pg/mL]", y="OD")

fit <- drm(formula = OD ~ logconc, data=melatonin.cal, fct=LL.4()
logDose=exp(10))

melatonin_curve.R

# (c) 2018 Viktoria Gastens

melatonin.time <- read.csv("ELISA_melatonin_2018-06-29_conc.csv", header=TRUE)
melatonin.time <- melatonin.time[grepl("*av", melatonin.time$Sample),]

p <- ggplot(melatonin.time, aes(x=ntime, y=Melatonin_pg_mL)) +
  geom_point() +
  geom_line() +
  geom_hline(yintercept=1.6, linetype="dashed") +
  scale_x_continuous(breaks=as.numeric(levels(as.factor(melatonin.time$ntime))),
  labels=times) +
  labs(x="Time of day", y="Melatonin concentration [pg/mL]"

ggsave("melatonin_time.png"
Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

Exploring links between mother and infant through human milk - Master's Thesis

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):

Gastens

First name(s):

Viktoria

With my signature I confirm that
- I have committed none of the forms of plagiarism described in the ‘Citation etiquette’ information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

Zurich, 28.07.2018

Signature(s)

Gastens

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.